

Psychological Review

THEODORE M. NEWCOMB, Editor
UNIVERSITY OF MICHIGAN

Lorraine Bouthilet, Managing Editor

CONTENTS

- The Relation of Response Latency and Speed
to the Intervening Variables and N in S-R
Theory.....KENNETH W. SPENCE 209
- Conditioning as an Artifact.....KENDON SMITH 217
- Do Intervening Variables Intervene?.....J. R. MAZE 226
- Response Factors in Human Learning.....GEORGE MANDLER 235
- Knowledge and Stimulus-Response
Psychology.....D. E. BERLYNE 245
- The Concept of Intelligence and the
Philosophy of Science.....CHARLES C. SPIKER
AND BOYD R. McCANDLESS 255
- The Skaggs-Robinson Hypothesis as an
Artifact of Response Definition.....MALCOLM L. RITCHIE 267
- Field Theory: II. Some Mathematical
Applications to Comparative Psychology.....WILLARD E. CALDWELL 271
- Critical Comment on "Learning and the
Principle of Inverse Probability".....ROBERT P. ABELSON 276
- A Metabolic Interpretation of Individual
Differences in Figural Aftereffects.....MICHAEL WERTHEIMER
AND NANCY WERTHEIMER 279

PUBLISHED BIMONTHLY BY THE
AMERICAN PSYCHOLOGICAL ASSOCIATION, INC.

CONSULTING EDITORS

SOLOMON ASCH
ROBERT BLAKE
STUART W. COOK
CLYDE COOMBS
LEON FESTINGER
W. R. GARNER
JAMES J. GIBSON
D. O. HEBB
HARRY HELSON
E. R. HILGARD
CARL I. HOVLAND
E. LOWELL KELLY
DAVID KRECH
ROBERT W. LEEPER

ROBERT B. MACLEOD
DAVID C. MCCLELLAND
G. A. MILLER
GARDNER MURPHY
OSCAR OESER
CARL PFAFFMANN
CARROLL C. PRATT
DAVID SHAKOW
RICHARD L. SOLOMON
ELIOT STELLAR
S. S. STEVENS
ERIC TRIST
EDWARD WALKER
ROBERT WHITE

The *Psychological Review* is devoted to theoretical articles of significance to any area of psychology. Except for occasional articles solicited by the Editor, manuscripts exceeding twelve printed pages (about 7,500 words) are not accepted. Ordinarily manuscripts which consist primarily of original reports of research should be submitted to other journals.

Because of the large number of manuscripts submitted, there is an inevitable publication lag of several months. Authors may avoid this delay if they are prepared to pay the costs of publishing their own articles; the appearance of articles by other contributors is not thereby delayed.

Tables, footnotes, and references should appear on separate pages; all of these, as well as the text, should be typed double-spaced throughout, in all manuscripts submitted. Manuscripts should be addressed to the Editor, Dr. Theodore M. Newcomb, Doctoral Program in Social Psychology, University of Michigan, Ann Arbor, Michigan.

PUBLISHED BIMONTHLY BY THE
AMERICAN PSYCHOLOGICAL ASSOCIATION, INC.
1333 SIXTEENTH ST. N. W., WASHINGTON 6, D. C.

\$6.50 volume

\$1.50 issue

Entered as second-class matter July 13, 1897, at the post-office at Lancaster, Pa., under Act of Congress of March 3, 1879

Acceptance for mailing at the special rate of postage provided for in paragraph (d-2), Section 34.40, P. L. & R. of 1948, authorized Jan. 8, 1948

Send all business communications, including address changes, to 1333 Sixteenth St. N.W., Washington 6, D. C. Address changes must arrive by the 25th of the month to take effect the following month. Undelivered copies resulting from address changes will not be replaced; subscribers should notify the post office that they will guarantee second-class forwarding postage. Other claims for undelivered copies must be made within four months of publication.

Copyright 1954 by the American Psychological Association, Inc.

THE PSYCHOLOGICAL REVIEW

THE RELATION OF RESPONSE LATENCY AND SPEED TO THE INTERVENING VARIABLES AND N IN S-R THEORY

KENNETH W. SPENCE

State University of Iowa

I

In the *Principles of Behavior* Hull introduced his theoretical constructs (intervening variables) initially in terms of independent environmental variables (e.g., S_e , N , \bar{I}_a , etc.), and completed the theoretical structure by anchoring them to certain response measures. The latter involved the introduction of specific *ad hoc* postulates that related each of the response measures (e.g., latency, frequency, resistance to extinction, etc.) to one or other of the theoretical constructs. Thus in the case of the response measure (latency) with which we shall be concerned, Hull made the following assumption:

The latency of response (R_t) is a decreasing hyperbolic function of the momentary effective excitatory potential (\bar{E}), i.e., $R_t = a\bar{E}^{-b}$, where a and b are empirical parameters (Postulate 13, p. 344).

Attention has been called elsewhere (7, 8) to the point that certain of these postulates are entirely superfluous in that assumptions already a part of the system (i.e., earlier postulates) permit one to derive a necessary relation between these response measures and one or other of the theoretical constructs. Thus in the case of the postulate introducing response probability or fre-

quency ($R\%$) as a function of effective excitatory potential (\bar{E}), the relation assumed is, as has been shown, derivable from definitions and postulates already made concerning effective excitatory potential (\bar{E}), oscillatory inhibition (O) and reaction threshold (L). In this particular instance, the postulate Hull made (normal integral function) happened to be identical with the relation that may be derived from assumptions already a part of the system (6). In the case of the latency measure, however, it may be shown that the postulate he assumed is actually inconsistent with a necessary relation that follows from earlier assumptions. The main purpose of the present article is to consider the implications for this relationship of the assumptions already made concerning \bar{E} , O , and L , and to extend their implications to the empirical laws (learning curves) to be expected between response latency and speed, on the one hand, and the variable N on the other.¹

We need to recall first that momentary effective excitatory potential, \bar{E} , is equal to $\bar{E} - O$, and that a response

¹ The assumptions made concerning O are those given in Hull's *Principles of Behavior* (3), not those in his later *Essentials of Behavior* (4) and *A Behavior System* (5).

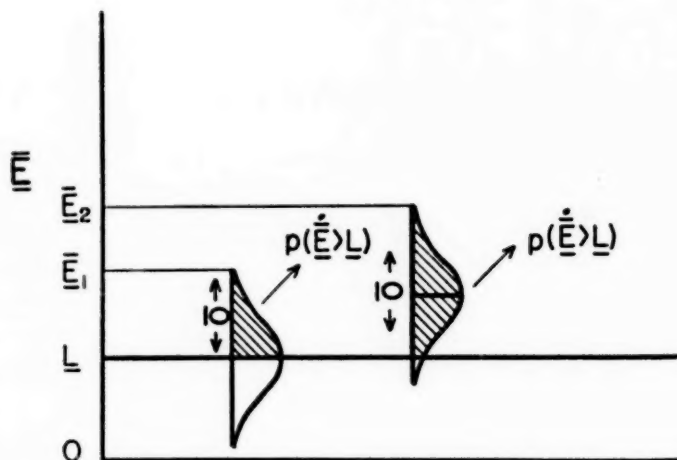


FIG. 1. The shaded portions of the upended normal distribution functions show the probability of \dot{E} being greater than L for two levels of \bar{E}

is assumed to occur to a stimulus only when (a) \bar{E} is greater than L , and (b) when an O value exists that is sufficiently small to make the value of \dot{E} greater than L . Oscillatory inhibition (O), it will be remembered, is assumed to change in value from moment to moment, the distribution of the values being postulated as normally distributed. The problem becomes one, then, of determining the average time, \bar{t} , before a momentary O value occurs that will provide an \dot{E} value greater than L .

Let P be the probability of occurrence of such an O value. Then the probability that such a value of O will be the first one to occur is P ; the probability that such a value will be the second is $(1-P)P$; the probability that it will be the third is $(1-P)(1-P)P$ or $P(1-P)^2$, etc. Considering now an indefinitely large number of occasions on which the stimulus is presented, and representing the number of occurrences of momentary O values on any occasion by n , we may weight each possible value of n by its probability (expected relative frequency), and thus obtain a

mean expected value of n . Estes (1) has shown that this mean value, \bar{n} , is equal to $1/P$. In other words, \bar{n} is the mean expected number of momentary O values that will occur on each stimulus occasion until an O value that provides a superthreshold \dot{E} value will occur.

If now we let u' represent the average time or duration of a momentary O , then the average time, \bar{t} , for a superthreshold \dot{E} value to occur will be the product of the expected number of momentary O values that will occur and the mean duration of a momentary O .

$$\bar{t} = \bar{n}u' = \frac{u'}{P}. \quad (1)$$

In terms of an average measure of speed of response evocation (\bar{v}), this equation becomes

$$\bar{v} = 1/\bar{t} = \frac{P}{u'} = uP, \quad (2)$$

where

$$u = 1/u'.$$

Figure 1 represents two levels of \bar{E} and shows their relation to L and O . The probability (P) of O being a value

that will produce a superthreshold \dot{E} , which can also be described as the probability that \dot{E} is superthreshold, $[p(\dot{E} > L)]$, is given by the proportion of the upended normal distribution that is above L . This yields

$$P = p(\dot{E} > L) = \int_{-\infty}^{\dot{E}-L-2.5\sigma_0} (O) dO. \quad (3)$$

Since \dot{E} is assumed to be an exponential function of N , i.e., $\dot{E} = A(1 - e^{-iN})$, it is possible to ascertain the theoretical function that P is of N by means of a table of cumulative probability values of the normal curve. Figure 2 shows the family of theoretical curves of the proportion of superthreshold \dot{E} values, $p(\dot{E} > L)$ as a function of N for different curves of growth of \dot{E} . When multiplied by the parameter, u , representing the reciprocal of average duration of a momentary O value, such curves provide a theoretical prediction of \bar{v} as a function of N . The relationship is identical in form, it

should be noted, with the theoretical frequency measure in classical conditioning (7) and implies an initial period of positive acceleration, providing that the data represent the major portion of the total possible learning and not just some later part of it.

Now the measure \bar{t} , it should be noted, is a measure of the time it takes to get the effector activity initiated. As such it does not involve the time or duration of whatever neuromuscular activity is involved on the part of the subject in the measuring situation. In actual practice, however, any measurement of response latency must also involve the time taken by the action of the effector system. Presumably our so-called latency measures, or measures of response time, represent a summation of these two durations so that, if we let T represent the obtained experimental measure of response time, \bar{t} the measure of action latency, and t' the duration of the effector activity involved in the meas-

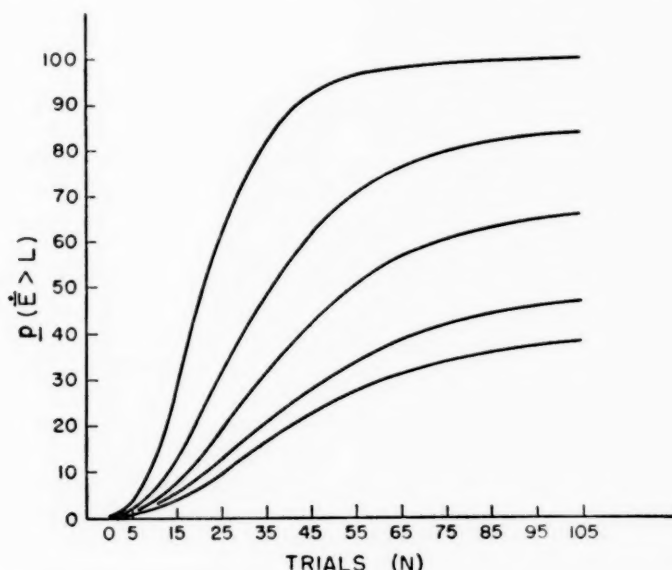


FIG. 2. Family of theoretical curves of the proportion of superthreshold \dot{E} values as a function of N for different curves of growth of \dot{E}

uring operation, then

$$T = \bar{t} + t'. \quad (4)$$

A similar equation can be derived for a measure of speed of response (V) as follows:

$$V = \frac{1}{\bar{t} + t'}. \quad (5)$$

The problem immediately arises as to how t' (and its reciprocal, v') vary with \bar{E} . The present writer has not been able to derive a relation between t' and \bar{E} from any of the existing postulates of the system. Accordingly, the working hypothesis is made that the relation is the simple hyperbolic one shown in the following equation:²

$$t' = \frac{c}{(\bar{E} - L)} \quad \text{where } c \text{ is a constant.}$$

Substituting now in equations 4 and 5, we obtain the following equations as representing the functions relating the experimental measurements of time (T) and speed of response (V) in simple instrumental conditioning to \bar{E} :

$$T = \frac{u'}{P} + \frac{c}{(\bar{E} - L)}, \quad (6)$$

$$V = \frac{1}{\frac{u'}{P} + \frac{c}{(\bar{E} - L)}}. \quad (7)$$

An interesting implication of the above theorizing is that the shape of the curve of V as a function of \bar{E} , and hence of N , will depend upon the magnitude of the parameter c , which is experimentally manipulable by varying the amount (duration) of motor activity involved in the measurement of the response. Thus in the simple approach type of situation (locomotion in a straight alley), one could vary the value of c by measuring the

response for different lengths of runway. When a minimum length of alley was employed, c would approach zero, and equation 7 would become identical with equation 2. Under this condition, the speed measure would provide an initially positively accelerated curve.

As the length of the runway involved in the measurement is increased, the value of c will increase, and the family of speed-of-response (v) curves as a function of N would be expected to vary in their initial phase from positive to negative acceleration. By selecting an appropriate length of runway, one should be able to obtain a curve that is linear in its early course of development. Finally, it should be noted that the theory implies that if the measurement (t') is taken from a point after the activity has started, the curve for speed of running (v') as a function of N should be a negatively accelerated exponential function, i.e.,

$$v' = \frac{A(1 - e^{-iN}) - L}{c}. \quad (8)$$

II

The above derivations concerning the relation of the measures, response time and speed of response to N in instrumental learning, assume that the growth of \bar{E} is an exponential one of the type that Hull employed. This assumption, in turn, depends upon the postulate that H grows in this manner and that the other factors determining \bar{E} , such as stimulus dynamism Q , drive D , incentive motivation K , and work inhibition I , are *constant* throughout the course of the training period. By means of various experimental techniques, such as distributing the trials, having only a few trials a day (3 or 4), etc., it is possible to keep Q , D , and I fairly constant. The situation is not so simple so far as the incentive moti-

² This implies, of course, that the relation between speed of activity (v') and \bar{E} is linear.

$$v' = \frac{(\bar{E} - L)}{c}.$$

vational factor, K , is concerned. In recent discussions of theories of learning (6) the writer has suggested that this factor K might represent a stimulus dynamism, that provided by the proprioceptive component (s_G) of the fractional anticipatory goal response (r_G). According to this notion, in instrumental learning involving reward, one also has classical conditioning taking place so that a fractional part of the goal response becomes conditioned to the stimulus situation. With conditioning, this r_G and its cue s_G , by virtue of generalization, move forward in time and thus become a part of the internal stimulus complex determining the strength of the instrumental response. Thus it is a kind of acquired motivating factor, the strength of which depends not only on the conditions of reinforcement (i.e., magnitude and delay of reinforcement), but also on the stage of training, i.e., number of reinforced trials (N).

This hypothesis has a number of important implications, not only for experiments involving different magnitudes and delays of reinforcement, but also for the nature of the learning (performance) curve in simple instrumental learning. Confining our interest here to the latter problem, we shall consider the implication of the variation of K during learning *without taking into account how this motivational variable interacts or combines with the other two motivational constructs, Q and D .*

According to our assumption concerning E , and ignoring D and Q ,

$$R = f(E) = (K \times H). \quad (9)$$

Substituting for K and H their postulated relations to N , we obtain the following:

$$R = f(E) = [B - (B - x)e^{-\rho N}] \times [A - (A - x)e^{-iN}]. \quad (10)$$

According to equation 9, the growth of E as a function of N would be initially positively accelerated instead of the negatively accelerated exponential function that would obtain if K were some constant value. In view of the fact that the curves of learning, e.g., $V = f(N)$ and $v' = f(N)$, are determined in part by the growth of E , it is readily apparent that we need to take this function into account in making any predictions as to the form of these curves. Thus, one interesting implication of this hypothesis is that if one were to establish K at a maximum (hence constant) value by setting up the classical CR ($S_c - r_G$) to the sight and sound of the lever (S_c) *prior to the beginning of the instrumental learning*, the growth of E would now be expected to be a negatively accelerated function throughout its course. Under this condition, the period of positive acceleration of the speed of response evocation curve $\bar{v} = f(N)$ should, other things being equated, be less than under the normal procedure in which the classical conditioning proceeds along with the instrumental learning.

In a similar fashion, one may predict that curves of speed of running ($v' = fN$) would be a negative growth function only under the condition in which K is a constant, and would tend to exhibit an initial phase of positive acceleration to the extent that K grows from zero to its maximum throughout the course of learning the instrumental response. It is probably the case that K is, as the result of transfer from past experience, already considerably developed in the simple running situation under normal illumination conditions, with the consequence that the curve of growth of E is distorted only slightly from the negative exponential function that holds when K is constant.

III

In this section some of the problems connected with the experimental testing of these theoretical implications will be discussed. If one examines the existing data from instrumental learning investigations, it will be found that a variety of curves of speed of response (reciprocal of latency or response duration) have been obtained. Of some score of curves, both from the literature and from unpublished studies conducted in the Iowa laboratory examined by the writer, it was evident that the majority showed a relatively brief initial period of positive acceleration followed by a period of prolonged negative acceleration to the asymptote. The second most frequently observed type was linear in its early portion (10 to 15 trials), followed by a negatively accelerated approach to the limit. A few curves exhibited negative acceleration throughout their course. It is the writer's impression that the latter type of curve tends to occur when the drive is very high, such as in running to escape an electric shock or under a strong hunger drive.

Unfortunately, these data are not very satisfactory for the reason that they are typically group curves involving the mean or median of a whole group of subjects. The forms taken by such group curves may deviate markedly from the curves of individual subjects, as Hayes (2) has recently so nicely demonstrated. The reasons for this become obvious when we consider that there are marked differences among individual subjects, not only in initial and final levels of performance, but also in their different rates of learning, i.e., their different rates of approach to the performance asymptote.

A number of alternatives to the use of such group curves suggest themselves. One is to employ the data of individual subjects. The notorious

variability of individual measures from trial to trial, however, usually necessitates some form of averaging of the individual measures in terms of blocks of trials, a procedure which also often leads to distortion of the curve. Thus if an individual curve that shows an initial phase of positive acceleration within the first 10 trials followed by a negatively accelerated phase is averaged in terms of successive blocks of 10 trials each, the initial period of positive acceleration will be lost entirely. The most satisfactory manner of treating individual measures, particularly measures of speed of response, would appear to be the moving average method with small blocks of trials, e.g., three trials in a block.

As an alternative to these curves based on individual measures and averages of a group of subjects, the writer has employed curves based on "like," or homogeneous, subjects. The homogeneity of the subjects can be ascertained in terms of the likeness of the subjects' performances at different stages of the practice or throughout the total learning period. Thus, in the case of the frequency measure in classical conditioning, the subjects' scores in terms of the total number of CR's occurring in a given number of trials, e.g., 100, can first be determined, and then groups of "like" subjects can be formed in terms of those that fall in a small range of scores, such as from 20 to 30 CR's, or from 50 to 60, etc. The writer has shown that such data from different parts of the distribution of total CR scores provide very smooth, comparable curves with relatively small numbers of subjects. In such curves the form is not a function of the distribution of the individual scores.

There are, of course, further refinements that can be made in this procedure. For example, in terms of the

speed measures we have been discussing, one could obtain measures for each subject at the beginning, at some intermediate point, and at the end of the learning, and then form groups of subjects that are alike at all three points rather than alike only on the basis of an over-all performance measure. As psychology moves into a period when testing of its theories requires the evaluation of the empirical data in terms of the precise form of some predicted lawful relation, more refined procedures for ascertaining the nature of the function will have to be made available.

A second point in connection with the testing of these theoretical implications is that the experimenter must define his response measures more precisely. Thus the present theory makes it necessary to differentiate between measures that involve starting of the action, the duration of the activity once set going, and combinations of both. The predicted differences in these various measures provide one of the most feasible ways of testing the theory. Similarly, the motivational conditions, relevant and irrelevant, will have to be carefully controlled and their possible differential effects taken into consideration.

Undoubtedly, the most difficult task will be that of arranging the experimental conditions so that there are no competing responses in the situation; for the theoretical model assumes but a single response, not a number of competing responses. Most instances of instrumental conditioning are really limiting cases of trial-and-error learning in which competing responses, while minimized, still play a more or less important role. Obviously, the occurrence of other competing responses in such simple learning situations involves time and thus importantly affects speed measures. More-

over, it is an unfortunate fact that interference from such competing responses is greatest at the beginning of instrumental learning when the rewarded response is weak. Later in the learning, this response is so much stronger than the competing ones that the latter are unable to interfere.

Elimination of competing responses will necessitate both experimental procedures and objective criteria for eliminating data on trials on which competing responses do occur despite the experimental controls. Particularly important also is control of the response orientation of the subject just prior to the presentation of the stimulus. A procedure that has been found to be quite successful in obtaining such control in a very few trials (two to four) is to employ two doors, one opaque and the other glass, in the starting box of an operant situation such as the straight alley. Each trial is started by the experimenter raising the opaque door first and then, at a fixed interval (three seconds) thereafter, the glass door. The raising of the opaque door serves as a warning signal for the subject, which very quickly comes to be set to respond to the raising of the glass door. Similarly, measuring techniques will have to be as precise as possible, particularly in measuring the speed of response evocation. The stop watch will hardly serve for our present purposes.

Finally, in addition to controlling competing responses, it will be necessary to minimize transfer from past experience so that the strength of the S-R is sufficiently low to provide a picture of the major portion of the total learning. It is particularly important from the point of view of the present theory that a good share of the initial phase of learning be represented in the data.

IV

In the preceding sections we have elaborated some of the implications with respect to the form of the speed of response evocation curve of learning of the original theoretical model put forward by Hull in his *Principles of Behavior*. The present treatment differs somewhat from that given by Hull in that certain implications of his postulates with respect to excitatory potential (\bar{E}), oscillatory inhibition (O), and threshold (L) were developed, with the consequence that it was possible to derive the relation to be expected between speed of response evocation (\bar{v}) and \bar{E} . By means of an assumption relating speed of running (v') to \bar{E} , additional implications were drawn concerning measures that involved various combinations of the two measures, speed of response evocation (\bar{v}) and speed of running (v'). Also considered were certain problems relating

to the obtainment and treatment of experimental data bearing on the theory.

REFERENCES

1. ESTES, W. K. Toward a statistical theory of learning. *Psychol. Rev.*, 1950, **57**, 94-107.
2. HAYES, K. J. The backward curve: a method for the study of learning. *Psychol. Rev.*, 1953, **60**, 269-276.
3. HULL, C. L. *Principles of behavior*. New York: Appleton-Century, 1943.
4. HULL, C. L. *Essentials of behavior*. New Haven: Yale Univer. Press, 1951.
5. HULL, C. L. *A behavior system*. New Haven: Yale Univer. Press, 1952.
6. SPENCE, K. W. Theories of learning. In C. P. Stone (Ed.), *Comparative psychology*. (3rd Ed.) New York: Prentice-Hall, 1951.
7. SPENCE, K. W. Mathematical formulations of learning phenomena. *J. exp. Psychol.*, 1952, **59**, 152-160.
8. SPENCE, K. W. *Symposium on relationships among learning theory, personality theory, and clinical research*. New York: Wiley, 1953.

(Received November 25, 1953)

CONDITIONING AS AN ARTIFACT

KENDON SMITH

*The Pennsylvania State University*¹

The case for a pure reinforcement theory of learning has been strongly put by recent papers (20, 21, 26, 40), and it is difficult now to escape the conviction that such a view is essentially correct.

In spite of its fundamental strength, however, reinforcement theory has remained weak in one respect. It has experienced continual difficulty in handling the problem of autonomic, visceral learning. To be consistent, Hull was obliged to maintain that reinforcement was crucial even in the presumed acquisition of such responses as salivation, alteration of skin resistance, cardiovascular changes, etc. This he did (19, pp. 76-80), although not as flatly as he is sometimes said to have done. Hull thus arrived at a position that many have regarded as untenable; for, although it is possible to imagine reinforcement factors at work in some instances of alleged visceral learning, there are many other instances in which the influence of such factors seems to be completely out of the question.

One variant of reinforcement theory that is designed to meet the situation more adequately is the so-called "two-factor" theory. It has been most recently and most vigorously promoted by Mowrer (29, 30), but Mowrer's views are in reality quite similar to those previously advanced by Thorndike (44, pp. 401-412) and Skinner (39, pp. 109-115). Mowrer differentiates between *conditioning* and *problem*

solving. The first term refers to the process of learning by a stimulus-substitution, contiguity principle; the second term, to that of learning by a principle of reinforcement. Conditioning is alleged to occur specifically via the autonomic nervous system, and problem solving to be mediated solely by the "central" (i.e., somatic) nervous system; a sharp distinction between the two neurological divisions is drawn and emphasized.

Such incisive dualism has an undeniable appeal. It suffers, however, from a disquieting lack of parsimony. It is extremely difficult to believe that "visceral learning" and somatic learning, so alike in so many respects, must be accomplished by two different principles (cf. 38). The resolution of the problem thus appears to lie elsewhere.

As it happens, there exists a rather simple way to save reinforcement, and parsimony with it. One can begin by accepting the notion that somatic learning is reinforcement learning. He can then go one step further than Mowrer has, and expunge from the viscera not only problem solving but conditioning as well. This procedure leaves the law of effect to rule in monistic majesty. Of course, it also makes the autonomic nervous system totally uneducable; but that, it can be asserted, is as it should be. For it can be argued that every "conditioned visceral response" is in reality an artifact, an innate accompaniment of the skeletal responses inculcated by the conditioning process.

The present paper will, in fact, attempt to defend this general line of thought.

¹ The writer wishes to acknowledge with real appreciation the helpful and sustained interest of his colleagues, Professors Joseph H. Grosslight, John F. Hall, and William M. Lepley.

THE ARGUMENT AND A SPECIFIC INSTANCE

In the elaboration of this proposal, it might be wise to begin with a concrete example. The galvanic skin response and its "conditioning" would seem to be appropriate.

When an attempt is made to condition the galvanic skin response, the procedure generally pairs a neutral conditioned stimulus with an unconditioned stimulus that is more or less noxious in nature, typically an electrical shock. Several pairings will eventuate, in some subjects, in a new correlation between the GSR and the originally neutral stimulus. "Conditioning"² has occurred.

Anyone who has actually carried out this procedure knows, however, that the foregoing account is seriously incomplete. The subject is not by any means a passive hulk during the entire program. In particular, he soon comes to regard the conditioned stimulus as a signal for a muscular "bracing" against the noxious stimulus to come; presumably this bracing, being somatic, is a matter of reinforcement learning. At any rate, one would expect such skeletal activity to be accompanied by the GSR, simply as a matter of innate neural connections. The occurrence of the conditioned GSR is thus hardly surprising, and it can be explained without recourse to a principle of autonomic learning. It is, in short, an artifact.

Now, at least two objections arise immediately. The first is that a scheme that works so well for one response, the

GSR, may not work at all for others: what about salivation, for instance? This is a reasonable misgiving, but it will be reserved for later consideration. The second objection is basic and serious, and it will be faced immediately.

The difficulty is this: Although it may seem quite natural to think of somatic responses as somehow "causing" their concomitant autonomic reactions, and although such a notion frequently finds expression in the literature, it is well to remember that the causal sequence is actually most obscure. It is entirely possible, for instance, that the muscular tension and its associated GSR arise as parallel events from a common innervation. The conservative view would seem to be that there is no distinguishable priority as between the two responses. By what right, then, is one considered the "real" response and the other a mere by-product?

Common sense would very likely answer that question rather readily: "The muscular response is the real response, because it is *conscious*; I didn't even know there was a 'galvanic skin response' until you told me about it!" Common sense thus suggests an answer which makes a certain amount of scientific sense too.

The combined, bracing-GSR reaction has already been pictured as a unitary one; and it was agreed to begin with that somatic responses attach themselves to new stimuli according to a reinforcement paradigm. It follows, therefore, that the combined response, as a unit, gets about by reinforcement. Now, the acquisition and maintenance of a voluntary (i.e., reinforcementally acquired) response is dependent upon the existence of afferent cues (cf. 4, pp. 524 ff.).³ The skeletal re-

² Inasmuch as this paper ultimately arrives at the conclusion that "conditioning" does not actually exist, the term itself has been treated so far in a rather gingerly fashion. In the interests of clean typography, quotation marks will be omitted from now on; but when conditioning and similar terms are used, they are meant to be read as if quotation marks were still present.

³ It is perhaps possible that this dependence stands as a testimony to the importance of immediate higher-order reinforcement; the af-

sponses, in the case at hand, provide a wealth of afferent information; but the autonomic reactions generate no regulatory feedback whatsoever. Acquisition of the whole response pattern, therefore, would seem to depend upon the integrity of the somatic component. If it were not for the muscular response, the GSR would not exist; but the GSR could be eliminated, perhaps by sympathectomy, and the bracing response would remain unaffected. The GSR is truly, in this sense, a secondary phenomenon, a by-product.

The earlier analysis of the conditioned GSR thus appears to be valid. At the same time, examination of this particular instance has generated a logic that can be applied to other instances of alleged autonomic learning. If it can be shown that the development of any new autonomic response is coincident with the growth of a somatic response known to have such an autonomic response as a regular correlate, it is legitimate to label the autonomic response as a secondary effect and the conditioning as bogus.⁴

The question remaining, then, is this: Do other instances of conditioning fit this pattern well enough to permit generalization of the artifact hypothesis?

APPLICATIONS IN OTHER INSTANCES

As one might expect, some varieties of conditioning conform to the pattern quite obviously, and others do not. For example, there are quite a few autonomic responses which, like the GSR,

ferent cues may constitute particularly prompt rewards and punishments. In any event, the dependence seems to be a fact.

⁴It is true that some visceral responses arouse afferent neural activity. As a class, however, these responses are sluggish, and the afferent information that they provide is greatly delayed in returning to the central nervous system. It can be presumed that there might as well be no return at all.

are known to be associated with the diffuse skeletal reactions that develop during a conditioning procedure. The foregoing discussion might just as well have revolved about the response of vasoconstriction as about the GSR; thus, conditioned vasoconstriction (2, 25, 37) fits the theory well. The same might be said for the few reported instances of conditioned vasodilation (2, 25), and for conditioned cardiac deceleration, which was recently observed in animals and human beings (22, 31). On the other hand, there are also instances that do not fall into place quite so neatly. Nevertheless (and this is the burden of the paragraphs to come), they do not, on close examination, present the prohibitive difficulties one might expect.

A case in point is that which generated the concept of conditioning in the first place: the salivary reflex. Everyone knows about Pavlov's classical experiments, and everyone knows that Pavlov induced his experimental animals to salivate on signal. It is not quite so widely known, however, that the animals did several other things besides salivating. Of special interest at the moment is the fact that they displayed gross skeletal responses. Pavlov observed movements of the head and "smacking . . . [of the] lips" (32, pp. 29-30), the animal appearing ". . . to take the air into its mouth, or to eat the sound" of the conditioned stimulus (33). Pavlov spoke frequently of the "alimentary reflex" or the "complex reflex of nutrition" (32, cf. pp. 13-14), and emphasized the fact that he was dealing with a pervasive pattern of behavior rather than with salivation alone.

These facts are manifestly made to order for the present hypothesis. Further, they have been confirmed by later experimenters. Zener, who also worked with the salivary reflex in dogs (48),

sometimes saw "chewing and licking" responses when the conditioned stimulus was administered; Zener and McCurdy (49) reported a correlation between rate of salivation and rate of chewing. And Moore and Marcuse, who stoutly undertook to condition the salivary responses of two sows (27), were explicitly concerned that the observed salivation might be due to oral activity. The records obtained by Moore and Marcuse seemed to give negative indications. These experimenters, nevertheless, had finally recognized the possibility, which had been so curiously neglected until their time, that the salivary responses were not essentially independent of the motor responses in the animal's behavior, but rather that the visceral activity of salivation was a natural concomitant of the acquired somatic activities of chewing, swallowing, etc. In this case, it might be noted, the possibility of concomitance had existed not only in the sense in which it has been developed in the foregoing discussion. It had existed also in the more obvious sense that sheer mechanical stimulation arising from oral activity could be expected to elicit salivation directly.

Of special interest at this point are Razran's reports of conditioned salivary responses in human subjects. Razran's experiments are well known, but it might be emphasized that, here once again, it was a complex process that was under investigation. Early results were rather widely irregular (34), and Razran soon came to recognize the importance of his subjects' "attitudes" (34, 35, 36). It turned out also that the act of thinking was important in the experimental results: "... it is seemingly not mere 'willing' but thinking of some more or less specific stimulus or response that produces a voluntary flow of saliva" (34, p. 9). Explicit instructions to "form associations" between conditioned and unconditioned

stimuli "produced very effective positive conditioning," while instructions to avoid forming associations usually had an opposite effect. It was even possible to substitute the thought of eating for actual eating, as the *unconditioned* stimulus, and still have a successful experiment.

Razran cautioned his subjects against gross oral movements, and his injunctions were evidently obeyed (34); thus the skeletal responses that might have evoked salivation were not as obvious in this instance as they were in the case of animal conditioning. Nevertheless, the fact that thought processes seemed so important is quite suggestive. If there is anything at all to a motor theory of thought and imagination, one must acknowledge the possibility of oral activity similar in nature, if not in magnitude, to that of the animal subjects. If such activity occurred, it might well have evoked the measured salivation. To be emphasized, also, is the role of gross bodily tension. Razran's subjects were hungry, and the food signal might well have led to a certain degree of general muscular relaxation in anticipation of ingestion. Such relaxation could be expected to tip the autonomic balance toward parasympathetic dominance and thus toward salivation.

It appears, then, that the conditioned salivary response is not beyond reconciliation with the notions proposed earlier. Attention may now be turned to another standard example of visceral learning: the conditioned pupillary response, which is of special interest in the present context principally because it evidently does not exist.

THE PUPILLARY RESPONSE

It is, perhaps, surprising to discover that pupillary conditioning is even in doubt. The early reports of Cason (5)

and Hudgins (17) were optimistic and well publicized. They mentioned not only iridic conditioning but even "voluntary control" (17) of the pupillary response. In 1934, Steckle and Renshaw reported an attempt to condition the pupillary reflex (42); the attempt was unsuccessful, and the experimenters expressed some skepticism about the earlier accounts. The ensuing discussion in the literature (18, 41) made it clear that these earlier studies had been replete with technical difficulties that left considerable room for subjective, judgmental factors. In 1936, in the second of the two papers last mentioned, Steckle again reported negative findings.

In 1938, positive results were claimed once more. Baker described iridic conditioning to "subliminal" and supraliminal sounds (1). Conditioning was alleged to be particularly rapid and stable when subliminal stimuli were employed. An elaborate repetition of Baker's work by Wedell, Taylor, and Skolnick (45) failed to confirm his results, as did a careful check by Hilgard, Miller, and Ohlson (16). A recent and thorough exploration by Hilgard, Dutton, and Helmick (14) has again produced very little evidence of successful iridic conditioning. Citing similar results with animals as well as with human beings (43), these latter authors warn that continued negative findings may force a revision of accepted learning theory.

In terms of the matter at hand, the defection of the pupillary response has a double significance. In the first place, it means that no theory need be seriously concerned with accounting for the existence of a conditioned iridic reflex; the present formulation, along with others, thus escapes a certain amount of travail. In the second place, and more importantly, the failure of

the pupillary response is materially embarrassing to a conditioning theory of visceral learning. Here, in the hands of competent workers, the principle of contiguity has had ample opportunity to exhibit itself. It has not done so. If "contiguity" and "artifact" are regarded as exhaustive alternatives, the artifact hypothesis appears to be the sole survivor.

It is of some incidental interest to speculate as to why conditioning should fail in the special instance of the pupillary response. It might be suggested that the changes in level of illumination that are customarily employed in iridic-conditioning experiments are of no practical consequence to the subjects, and that, therefore, no anticipatory skeletal responses are acquired. There being no acquired skeletal responses, there is no conditioning.

In this connection, it is worth noting that unexplained failures characterize almost every conditioning experiment. Some subjects condition, and others do not. Such vagaries are much more suggestive of individual differences in personality structure than they are of the presumed essential similarity of fundamental neurological processes from person to person. One thinks of greater and lesser tendencies toward anticipation, anxiety, and (literal) tension as perhaps underlying the personal variations in conditionability. It is evident that the recently reported correlations between anxiety level and conditionability fit in rather neatly with such a conceptualization (3, 31, 46).

TESTS AND TESTABILITY

The foregoing evidence is taken to be strongly favorable to an artifact theory of conditioning. It hardly suffices to close the matter, however. It seems appropriate now to examine other lines of evidence, formal and informal, actual

and potential, which might have a bearing upon the hypothesis at hand.

From everyday experience comes the first datum: it is notoriously possible to forestall an untoward visceral response by "not thinking about it." Stimuli that might otherwise elicit nausea, flushing, or sexual reflexes lose much of their effectiveness when one refuses to dwell on their implications. If thought is somatic, it would thus appear that the autonomic response depends upon the somatic and does not arise independently of it. There are clear and perhaps important implications here for a further understanding of such phenomena as hysterical nausea and psychological impotence.

A somewhat similar argument also arises from everyday experience. Sheffield (38) has well pointed out the obvious fact that bowel and bladder functions are in some sense subject to reinforcement training. An equally obvious fact, to be emphasized here, is that such training is conducted in terms of massive somatic responses, and that everyday visceral control is exercised only by virtue of diffuse contractions of the skeletal musculature. Again, it appears that the "real" response is a somatic one.

Both of the preceding arguments bolster the basic contention that learned visceral activity is a by-product of somatic behavior. A crucial test of this contention could be made if one could provide oneself with a subject who is completely passive, and even unthinking. A subject in such a state should not, according to theory, be susceptible to conditioning. Could such a situation be contrived under deep hypnosis? One can imagine a good hypnotic subject, instructed to relax and to "keep his mind completely blank," but to remain awake and aware of such stimuli as bells and shocks; would the conditioned GSR appear in such circumstances?

To the best of the writer's knowledge, such an experiment has not been performed. It must be admitted, however, that a rather similar situation has been reported, and that it does not seem to conform to theoretical expectations. In 1938, Lindsley and Sassaman (24) reported the discovery of an individual who was able to exercise "voluntary control" over his pilomotor responses. The body hair could be erected or lowered upon signal. The subject reported having come upon his talent more or less accidentally. He denied that the basis for his performance was controlled imagery: he imagined nothing in particular; he simply "willed" the response. Neither were there obvious skeletal movements or muscular tensions to account for these hirsute accomplishments. This case, anomalous as it is, presents difficulties for practically any theory of conditioning. The only saving feature, as far as the present hypothesis is concerned, is that there were indications of a very diffuse activity of some sort: cardiac and pupillary effects accompanied the pilomotor responses. The possibility of a covert somatic response was thus considerable.

The tenor of this discussion suggests a well-known animal investigation, that of Harlow and Stagner (12). Using cats as subjects, these investigators paralyzed the striate musculature by deep curarization. While the animals were under curare, the pupillary response to electrical shock (which could still be elicited) was conditioned to the sound of a buzzer. Here again is ostensibly negative evidence. As it has turned out, however, the effects of curare upon the neuromuscular system seem to be quite complex, and later experiments have found striate-muscle responses even in deeply curarized animals (10, 11). If Harlow and Stagner's animals retained some degree of skele-

tal responsiveness, the experiment loses much of its crucial aspect as far as the present discussion is concerned.

One final bit of information comes from Mowrer. In a recent defense of two-factor theory (30), he has quoted incidental observations by Gantt to the effect that when a conditioned-response pattern embraces both motor and cardiac elements, the visceral, cardiac component often persists after the motor has disappeared. Such an observation, if well founded, might also damage an artifact theory. The original passage from Gantt goes on to say, however, that "the respiratory component also accompanies the cardiac" (9, p. 51). Evidently, then, Gantt means by "motor component" only really gross skeletal behavior. Less, spectacular somatic responses not only could occur along with the autonomic responses, but admittedly do.

It begins to appear that an artifact theory is remarkably easy to defend. As a matter of fact, it is almost invulnerable. Unless an experimenter is meticulous in the elimination of skeletal responses, it will always be possible to account for visceral conditioning by postulating undetected somatic activity. And whenever visceral conditioning fails, one can always claim that no effective skeletal behavior was aroused (a stand already taken above with respect to the failure of pupillary conditioning).

This is an unfortunate situation scientifically, but it is not unique. In the current controversy over "latent learning," for instance, the reinforcementists find themselves in exactly the same logical position. The strategy in that instance has been to attempt to eliminate all conceivable opportunities for reinforcement to operate, expecting thus to minimize "latent learning." Perhaps an analogous attack can be made upon the problem here at hand.

RELATED NOTIONS

The concepts sketched above are not, of course, completely original. There has long existed a general feeling that the conditioned reflex is somehow vaguely fraudulent. Hilgard, who has displayed a lasting interest in the systematic status of conditioning (cf. 13, 14, 15, 16), has quite recently taken a position essentially antecedent to the one adopted here. Two excerpts from his 1948 book will define his stand:

... experiments in which autonomic conditioning takes place (salivation, galvanic response) are full of indirect accompaniments [of a reinforcement variety]. When the circumstances seem almost ideal for demonstrating ... conditioning, as in attempts to condition pupillary contraction by presenting a tone along with a light, it is extremely difficult to obtain any conditioning at all. The few cases which have found conditioning are in doubt ... (13, pp. 119-120).

... there is little evidence that the simultaneous or nearly simultaneous occurrence of an incidental stimulus and an unconditioned response is the sufficient condition for establishing a sensori-motor association between them (13, p. 334).

An earlier (1935) and more general expression of somewhat the same sentiment can be found in a note by Foley:

... experimental work on conditioning in complex organisms ... is often completely vitiated by the fact that certain implicit reactions, with their attendant stimulations, are frequently occurring simultaneously or temporally adjacent to the predetermined stimulations experimentally administered by the investigator (7, p. 444).

Foley, however, does not fix his suspicions upon reinforcement factors as plainly as does Hilgard. Neither did Freeman, who summarized his experiments of 1930, on the GSR, in much the same vein:

Since the quantitative results were obtained in a manner similar to that of animal experimentation, I have interpreted them as "conditioned responses"; but it seems that they might just as easily be interpreted as physio-

logical accompaniments of the attitudes of expectation and surprise (8, p. 534).

Cook and Harris (6), Lazarus and McCleary (23), and Mowrer in an earlier paper (28) have expressed similar opinions on the conditioned galvanic skin response. With respect to salivary conditioning, Razran (cf. 34, 35, 36) and Zener (48) have also emphasized motivational and attitudinal factors in quite general terms.

The formulation arrived at in the present paper should not be confused with an already rather common one. A great many theorists have maintained, in one way or another, that what actually develops as a result of stimulus contiguity is an association, a cognition, an expectancy, or, perhaps, a conditioned sensation. At any rate, this general view represents essentially an extension of the contiguity principle to somatic learning (it is not always completely clear how these theorists account for the autonomic phenomena that arise during the conditioning process). Such an extension of the contiguity principle is, of course, quite opposed to the theory at hand, which sets out to obliterate the principle entirely.

The present view welcomes the suggestion that what really happens in visceral conditioning is that an "expectancy" develops. It would insist, however, that such an "expectancy" is a somatic process, inculcated by reinforcement learning (cf. 47; 13, pp. 331-334) and occurring in the skeletal musculature, and that, as an innate by-product of this muscular activity, the observed visceral responses arise.

REFERENCES

1. BAKER, L. E. The pupillary response conditioned to subliminal auditory stimuli. *Psychol. Monogr.*, 1938, **50**, No. 3 (Whole No. 223).
2. BEIER, D. C. Conditioned cardiovascular responses and suggestions for the treatment of cardiac neuroses. *J. exp. Psychol.*, 1940, **26**, 311-321.
3. BITTERMAN, M. E., & HOLTZMAN, W. H. Conditioning and extinction of the galvanic skin response as a function of anxiety. *J. abnorm. soc. Psychol.*, 1952, **47**, 615-623.
4. BORING, E. G. *Sensation and perception in the history of experimental psychology*. New York: D. Appleton-Century, 1942.
5. CASON, H. Conditioned pupillary reactions. *J. exp. Psychol.*, 1922, **5**, 108-146.
6. COOK, S. W., & HARRIS, R. E. The verbal conditioning of the galvanic skin reflex. *J. exp. Psychol.*, 1937, **21**, 202-210.
7. FOLEY, J. P. A critical note on certain experimental work on the conditioned response. *J. gen. Psychol.*, 1935, **12**, 443-445.
8. FREEMAN, G. L. The galvanic phenomenon and conditioned responses. *J. gen. Psychol.*, 1930, **3**, 529-539.
9. GANTT, W. H. Psychosexuality in animals. In P. H. Hoch & J. Zubin (Eds.), *Psychosexual development in health and disease*. New York: Grune & Stratton, 1949. Pp. 33-51.
10. GIRDEN, E. Generalized conditioned responses under curare and erythroidine. *J. exp. Psychol.*, 1942, **31**, 105-119.
11. GIRDEN, E., & CULLER, E. Conditioned responses in curarized striate muscle in dogs. *J. comp. Psychol.*, 1937, **23**, 261-274.
12. HARLOW, H. F., & STAGNER, R. Effect of complete striate muscle paralysis upon the learning process. *J. exp. Psychol.*, 1933, **16**, 283-294.
13. HILGARD, E. R. *Theories of learning*. New York: Appleton-Century-Crofts, 1948.
14. HILGARD, E. R., DUTTON, C. E., & HELMICK, J. S. Attempted pupillary conditioning at four stimulus intervals. *J. exp. Psychol.*, 1949, **39**, 683-689.
15. HILGARD, E. R., & MARQUIS, D. G. *Conditioning and learning*. New York: D. Appleton-Century, 1940.
16. HILGARD, E. R., MILLER, J., & OHLSON, J. A. Three attempts to secure pupillary conditioning to auditory stimuli near the absolute threshold. *J. exp. Psychol.*, 1941, **29**, 89-103.
17. HUDGINS, C. V. Conditioning and the voluntary control of the pupillary light reflex. *J. gen. Psychol.*, 1933, **8**, 3-51.

18. HUDGINS, C. V. Steckle and Renshaw on the conditioned iridic reflex: a discussion. *J. gen. Psychol.*, 1935, **12**, 208-214.
19. HULL, C. L. *Principles of behavior*. New York: D. Appleton-Century, 1943.
20. KENDLER, H. H. Reflections and confessions of a reinforcement theorist. *Psychol. Rev.*, 1951, **58**, 368-374.
21. KENDLER, H. H., & UNDERWOOD, B. The role of reward in conditioning theory. *Psychol. Rev.*, 1948, **55**, 209-215.
22. KOSUPKIN, J. M., & OLMSTEAD, J. M. D. Slowing of the heart as a conditioned reflex in the rabbit. *Amer. J. Physiol.*, 1943, **139**, 550-552.
23. LAZARUS, R. S., & McCLEARY, R. A. Autonomic discrimination without awareness: a study of subception. *Psychol. Rev.*, 1951, **58**, 113-122.
24. LINDSLEY, D. B., & SASSAMAN, W. H. Autonomic activity and brain potentials associated with "voluntary" control of the pilomotor (mm. arrectores pilorum). *J. Neurophysiol.*, 1938, **1**, 342-349.
25. MENZIES, R. Conditioned vasomotor responses in human subjects. *J. Psychol.*, 1937, **4**, 75-120.
26. MILLER, N. E. Comments on multiple-process conceptions of learning. *Psychol. Rev.*, 1951, **58**, 375-381.
27. MOORE, A. U., & MARCUSE, F. L. Salivary, cardiac and motor indices of conditioning in two sows. *J. comp. Psychol.*, 1945, **38**, 1-16.
28. MOWRER, O. H. Preparatory set (expectancy)—a determinant in motivation and learning. *Psychol. Rev.*, 1938, **45**, 62-91.
29. MOWRER, O. H. On the dual nature of learning—a re-interpretation of "conditioning" and "problem solving." *Harv. educ. Rev.*, 1947, **17**, 102-148. (Reprinted in Mowrer, O. H., *Learning theory and personality dynamics*. New York: Ronald, 1950. Pp. 222-274.)
30. MOWRER, O. H. Two-factor learning theory: summary and comment. *Psychol. Rev.*, 1951, **58**, 350-354.
31. NOTTERMAN, J. M., SCHOENFELD, W. N., & BERSH, P. J. Conditioned heart rate response in human beings during experimental anxiety. *J. comp. physiol. Psychol.*, 1952, **45**, 1-8.
32. PAVLOV, I. P. *Conditioned reflexes*. London: Oxford Univer. Press, 1927.
33. PAVLOV, I. P. The reply of a physiologist to psychologists. *Psychol. Rev.*, 1932, **39**, 91-127.
34. RAZRAN, G. H. S. Conditioned responses: an experimental study and a theoretical analysis. *Arch. Psychol.*, 1935, No. 191.
35. RAZRAN, G. H. S. Attitudinal control of human conditioning. *J. Psychol.*, 1936, **2**, 327-337.
36. RAZRAN, G. H. S. Conditioning and attitudes. *J. exp. Psychol.*, 1939, **24**, 215-226.
37. ROESSLER, R. L., & BROGDEN, W. J. Conditioned differentiation of vasoconstriction to subvocal stimuli. *Amer. J. Psychol.*, 1943, **56**, 78-86.
38. SHEFFIELD, F. D. The contiguity principle in learning theory. *Psychol. Rev.*, 1951, **58**, 362-367.
39. SKINNER, B. F. *The behavior of organisms*. New York: D. Appleton-Century, 1938.
40. SPENCE, K. W. Cognitive vs. stimulus-response theories of learning. *Psychol. Rev.*, 1950, **57**, 159-172.
41. STECKLE, L. C. Two additional attempts to condition the pupillary reflex. *J. gen. Psychol.*, 1936, **15**, 369-377.
42. STECKLE, L. C., & RENSHAW, S. An investigation of the conditioned iridic reflex. *J. gen. Psychol.*, 1934, **11**, 3-23.
43. STERN, F. An investigation of pupillary conditioning. Unpublished doctor's dissertation, Univer. of Washington, 1948.
44. THORNDIKE, E. L., et al. *The fundamentals of learning*. New York: Teachers Coll., Columbia Univer., 1932.
45. WEDELL, C. H., TAYLOR, F. V., & SKOLNICK, A. An attempt to condition the pupillary response. *J. exp. Psychol.*, 1940, **27**, 517-531.
46. WELCH, L., & KUBIS, J. The effect of anxiety on the conditioning rate and stability of the PGR. *J. Psychol.*, 1947, **23**, 83-91.
47. WOODWORTH, R. S. Reinforcement of perception. *Amer. J. Psychol.*, 1947, **60**, 119-124.
48. ZENER, K. The significance of behavior accompanying conditioned salivary secretion for theories of the conditioned response. *Amer. J. Psychol.*, 1937, **50**, 384-403.
49. ZENER, K., & MCCURDY, H. G. An analysis of motivational factors in conditioned behavior: I. The differential effect of changes in hunger upon conditioned, unconditioned, and spontaneous salivary secretion. *J. Psychol.*, 1939, **8**, 321-350.

(Received August 6, 1953)

DO INTERVENING VARIABLES INTERVENE?

J. R. MAZE

University of Sydney

There is a kind of fallacy to which psychology seems especially prone (although it is not by any means peculiar to it) and which has been pointed out many times under different names—e.g., “faculty-naming,” “hypostatization,” “the postulation of imaginary forces,” “verbal magic.” But its logical structure has rarely been made explicit, and one finds that even those writers who attack it frequently commit it themselves. It seemed on first reading that MacCorquodale and Meehl, in their very provocative and valuable paper (11), had gone a good way toward clarifying this problem, but the fact that their authority has been invoked in support of such widely diverse views, including methodologies that appear particularly disposed toward hypostatization, might be taken as a sign that their paper has failed of its effect, and that a re-examination of it might be profitable.

Its main burden was to draw a distinction between two kinds of theoretical concept—“intervening variables,” whose meaning and truth were completely reducible to those of the “empirical relationships” with which they dealt, or which they described, and “hypothetical constructs,” which had “surplus meaning” so that the truth of statements involving them was *not* completely reducible to the truth of statements about the empirical relationships in connection with which they were hypothesized. MacCorquodale and Meehl do not contend that either of these kinds is illegitimate. What they do object to is the surreptitious use of intervening variables as if they were hypothetical constructs—as if they could

sustain the functions of the latter. So far, the argument seems perfectly sound, and the fallacy to which they are pointing seems to be the “verbal magic” one, i.e., giving a name to a certain kind of event and then using that name as if it accounted for the *occurrence* of that kind of event. One example of this, I suggest, might be to use the phrase “having a valence” as meaning “being the object of our striving,” and then seeming to account for our striving for something by saying that it has a valence for us.

But when one considers their criticism of “libido,” one feels that they do not quite make the point that the situation one cannot use the intervening variable to account for is just the situation whose name it is. They simply say that “certain puzzling phenomena are *deduced* (“explained”) by means of the various properties of libido . . .” (11, p. 105). And in their examples of “pure” intervening variables, we find some that could hardly avoid being used in the illegitimate way—e.g., “valence” itself—and some that seem quite distinct from the sort of “calculational device” (to use Spence’s phrase) which, following Tolman’s scheme for “breaking down into more manageable form the original complete f_1 function” between behavior and the independent experimental variables (17), they originally offer as intervening variables.

This confusion, I contend, comes about mainly because they are not clear on what the “empirical relationships” they are discussing are between, or, in general, what sort of thing can have relations. Thus, by manipulating the four variables that enter into Hull’s

equation for habit strength (9)—number of reinforcements, delay in reinforcement, amount of reward, and stimulus-response asynchronism (N , t , w , and t')—they show clearly that intervening variables in the “calculational device” sense depend on quite arbitrary groupings of the empirical variables concerned, and contend that from these four they “could define 15 alternative and equivalent sets of intervening variables” (11, p. 98). But concerning one of these, a new intervening variable involving only N , and which they suggest might be called “cumulative reinforcement,” they say: “Suppose now that a critic asks us whether our ‘cumulative reinforcement’ really *exists*. This amounts to asking whether we have formulated a ‘correct statement’ concerning the relation of this intervening variable to the anchoring (empirical) variables.” Now, to ask whether *its* relation to the empirical variables is correctly stated is already to treat it as “existing,” although this question of existence, which is brought in at many points throughout the essay, is not relevant since, as Bergmann (3) points out, ratios between the quantities of different variables exist even though the number by which a ratio may be expressed may not itself stand for a quantity of anything. More importantly, to ask about “its” relation to anything is to treat it as being the sort of thing that can be a term of a relation—that is, as being qualitative, as being some state or condition or “stuff” that there can be quantities of. Such a notion must always have “surplus meaning”; that is, a term of a relation must have some nature, some collection of properties, other than its having that relation; otherwise it would be unintelligible to say that *it* had that relation. What would “it” refer to?

Now, such a conclusion is precisely what MacCorquodale and Meehl want

to avoid, but it is entailed by their speaking of “its relation to the empirical variables.” What should really be considered is whether a “correct statement” of the relationship of the *number of reinforcements to response strength* has been formulated. Concerning sH_R in general, the point is that its “existence” is a question of whether or not the functional relationship between environmental variables and response strength has been correctly stated, which is precisely why Hull retained the S and the R in his notation: that is, in order to stress the point that it is a mathematical relationship between something done to the organism and something done by the organism, and is not in itself in any precise sense a description of the organism.

It is interesting, however, to note that Hull himself makes precisely the same error in giving his developed account (10) of what he means by the apocalyptic phrase “anchoring variables at both ends.” He says: “. . . my own system . . . requires that the habit strength (sH_R), afferent impulse (s), and drive intensity (D) must each be calculable from their antecedent conditions, that the nature and magnitude of the reaction potential (sE_R) must be calculable from the values of sH_R , s , and D taken jointly, and that the nature and magnitude of the several reaction functions . . . must each be calculable from sE_R ” (10, p. 281). But the point is that the formulae for calculating “habit strength,” “drive,” etc. from “their” antecedent conditions were derived in the first place from the series of experiments in which variations in each of the environmental variables listed (the others being held constant) were correlated with variations in the different measures of response strength, not in any sense with variations in “intervening variables.” Thus this process

of calculation and verification that Hull is describing is just a matter of verifying those empirical correlations with response strength for *different values* of the antecedent conditions—i.e., of checking the curve fitting at different points of the curve.

We have to be clear, then, that the empirically found mathematical relations are not between (for example) sH_R and its antecedents on the one hand, and between sH_R and its consequents on the other, but just between the antecedent and consequent events—in fact, that that relationship is just what sH_R is. The other notion of what it is—i.e., some qualitative condition of the organism, a given amount of it being produced by specified environmental events, and in its turn producing a given strength or frequency of response—is what MacCorquodale and Meehl refer to as a hypothetical construct; but, as I suggested, their description of the meaning of intervening variables seems in some places to imply that sort of notion.

Now that might seem hardly more than a slip of the pen on their part, especially since they go on to say (11, p. 99) that “when habit strength *means* the product of the four functions of w , t , t' , and N , then if the response strength is related to these empirical variables in the way described, habit strength ‘exists’ in the trivial sense that the law holds”—although one cannot see why this should be thought trivial if nothing else was ever expected of sH_R . But if that really is what MacCorquodale and Meehl mean by “intervening variable” then why do they not press home their sketched-in criticism of any rigid hierarchy of intervening variables? They suggest that Tolman’s argument (17) that it is easier to determine the complicated f_1 function by parts than “as a whole” is “not very cogent.” They do not elaborate this point, but

I take it that what they are suggesting is that one can only do it step by step, so that setting up a doctrine which advocates the use of such intermediate steps cannot be any contribution (if that is really all it offers). One might say that if we want to discover the product of 3, 4, and 5, then it is not really possible to multiply them all together at once; we have to find the product of, say, 3 and 4 and then multiply that product by 5. But that procedure has no advantage whatever over first taking 3 and 5 together, or 4 and 5, so long as we really are concerned only with convenience of computation. In their manipulation of w , t , t' , and N , MacCorquodale and Meehl provide all the materials necessary for showing that the same applies to the grouping of *all* the environmental variables studied by Hull, including as well as those four the maintenance and stimulus variables effective at the time of any given trial. If convenience of manipulation is the only concern, it must be very difficult to show why that particular four out of all these factors should be taken together, and why their product should be given a special name.

Hull, too, feels sensitive on this point, since he puts a similar objection into the mouth of “his friend Woodrow” (10, p. 284). His replies to it are quite unconvincing: he says in the first place that even if we did put all the subordinate equations together to form one, it would be found still to contain “in some form or other the mathematical equivalents of the various equations linking the observable and the hypothetical unobservable elements of the situation.” But to say that the groupings are merely for convenience is to deny that one is thinking of any “hypothetical unobservable elements,” and as for the rest, to say that the mathematical representatives of the *observable* elements (and the relations between

them) would still be found in the master equation is to admit the point that "Woodrow" is making.

Hull's second point in reply to the criticism gives up the attempt to justify intervening variables on a formal basis and introduces the material consideration that the action of the four training variables might be temporally remote from that of the critical stimulus and the state of need. And "while it is perfectly possible to put into a single equation the values of events which occur at very different times," still those past events cannot be causally active now, and " $\dot{s}H_R$ is merely a quantitative representation of the perseverative after-effects" of those four training variables (10, p. 285). Hull, then, groups the variables in this regular way not merely as a matter of convenience in calculation (if indeed one way could be more convenient than another), but also because he regards the variables in any group as acting together to build up some specific condition in the animal, and the intervening variable based on that group would then be thought of, if not actually as a "measure" of that condition, at least as varying quantitatively in direct relation with it.

One might contend, then, that some such speculation about the accompanying qualitative states could be the only reason for clinging to a set order of groups and giving names to their products when these products have still to be combined to discover the probability of a given response. But as "calculational devices," the only solid content for the notion of "intervening variables" is just the collection of correlation coefficients *between* response strength and each of the empirical variables found relevant to it. Without some such mathematical form we could not give even this slight meaning to the term "intervening," since we would be left only with the assertion that a given

factor has some causal connection with a given response, and causality does not in any sense *intervene between* antecedents and consequents—it is just the antecedent *event* that produces the consequent, not "causality."

The finding of such relationships, by the way, indicates the only sound meaning for the phrases "empirical variables" and "empirical relationships"—that is, what is really meant is more like "empirically-found relationships," and even here, for the empiricist, the phrase is redundant, since there is no other way of arriving at knowledge than by finding it empirically. By using the word in this way, MacCorquodale and Meehl convey a vague suggestion that empirical relationships are to be compared, unfavorably, with more certain, more fundamental, more intelligible "laws," and it is strange that this hint of the rational is to be found amongst the positivists at large (e.g., 7).

One point that helps to preserve this distinction in thought is that it is very difficult to discover perfect regularity in the sense of coextension when we are looking for causal connections. Frequently one can find only conditions that are sufficient but not necessary, or necessary but not sufficient. This has also led to the current suspicion that there must be something wrong with causality, has led to the not-talking about it, and to the being content with (and compensatory exaltation of) correlation coefficients and statements of probability in general. Without denying the usefulness of such mathematical procedures as a first approach to a confused field, one can still say that the rejection of indeterminism will involve denying the adequacy of probabilities in science, and will involve affirming the presence of a criterion for every case in which a necessary condition is not also sufficient, and vice versa. That is to say, it is always theoretically pos-

sible to extend our knowledge of the effective conditions until we arrive at a set that is necessary *and* sufficient for any given event. In seeking causes, such criteria can only be found by considering *both* the nature of the thing in which the change is produced and the nature of the thing acting on it (1). A recognition of this principle seems characteristic of the work of those scientists who increase the number of *general* propositions known in their subjects. To take only two examples in psychology, there is the recent work of Tinbergen (16) with his emphasis on the need for both a specific type of stimulus situation and a specific bodily condition for the production of an instinctive response; and there are the explicit formulations of Freud (8), which might have prevented a good deal of the dreary heredity-environment controversy if they had been better known. Watsonian behaviorism, then, was based on a false premise—"... given the response the stimuli can be predicted; given the stimuli the response can be predicted" (18, p. 167)—since the same stimulus situation will produce different responses in the same animal according to changing conditions in the animal, a point recognized in part by Hull in his emphasis on *D*.

If psychology is ever to make predictions, then, rather than mere statements of probability, it is, in MacCorquodale and Meehl's terms, committed to "hypothetical constructs" or, more precisely, to making hypotheses (and trying to verify those hypotheses) about what processes in the animal mediate a given change in its behavior. As Bergmann points out (4), these hypotheses will be assertions that processes of a kind which we have known in other places are going on (so far unobserved) in this place—the point being that we can arrive at the notion of any term or kind only by confronting such a kind,

only from experience. Where else could we get the material for our fantasies? Wherever we seem to "construct" the notion of a kind, it is always by conjoining properties that we have encountered (separately) in actual things—and if that is so, then it is sufficient to dispose of the doctrine that there *can be* "convenient fictions" in science, when these are said to be neither true nor false. MacCorquodale and Meehl themselves reject this latter doctrine, but their retention of the phrase "hypothetical constructs" renders them liable to be misunderstood as supporting it (cf. Bergmann, 3), as does their question whether constructs are "existential"—as though it were a real possibility that we could talk about some *thing* which was not existential.

Now, "intervening variables" in the sense described above—as correlations between events impinging on the organism and responses produced by the organism—cannot be states or properties or qualities of the organism itself, even though they presuppose the existence of such qualities. The environmental variables may have been grouped in a specific way (as Hull's are) because of some tentative speculation about a specific change produced in the animal by each group, but even so, and even if these speculations are put forward, the mathematical relations remain distinct from the hypothesized processes which are held to account for them. This is a point overlooked by MacCorquodale and Meehl when they say (11, p. 101) that "there are various places in Hull's *Principles* where the verbal accompaniment of a concept, which in its mathematical form is an intervening variable in the strict (Tolman) sense, makes it a hypothetical construct." That makes it seem that they always have half thought of intervening variables as being in some very vague sense "in" or "of" the organism (and have accepted

them thus as legitimate), the only point being that they are not to be ascribed properties, not to be characterized in any way (cf. O'Neil, 12), else they become hypothetical constructs. This suggestion is borne out in their discussion of Skinner's treatment of emotion as "a 'state of the organism' which alters the proportionality between reserve and strength." They say (p. 102): "The 'state' of emotion is not to be described in any way except by specifying (a) The class of stimuli which are able to produce it and (b) The effects upon response strength. Hence emotion for Skinner is a true intervening variable, in Tolman's original sense." Now, previously they had described "state" as a "wholly noncommittal word" which specifies "nothing except that the conditions are internal" (p. 97). But this point about internality is precisely the one at issue in discussing the fallacy of hypostatization. We may agree that there always will be some condition of the organism in virtue of which, in specified circumstances, a given response will be produced, but, in direct contradiction to MacCorquodale and Meehl, it *must* be described in ways other than its relations to its antecedents and its consequences (or in general its relations to anything); otherwise (*vide supra*) statements about what produced *it* and what *it* produced (or about any of its relationships) will not be intelligible. In saying that it is not to be described except by specifying those relations, and in saying (by calling it a "true intervening variable") that it is identical with the "empirical relations" and yet is a state of the organism, MacCorquodale and Meehl are setting up the notion of something whose whole nature it is to stand in a given relationship. Even if it is objected that it is the organism that has the relationships, and that they are not its whole nature (since it has many

other properties) but only a part of it, still this modified doctrine faces the same difficulty—namely, that it is strictly "unspeakable" (2) since we can only grasp a relationship if we can distinguish the terms that have it to each other; that is, seeing them as distinct, as having distinct natures, is a part of seeing them as related. If we say that its relationship to a certain stimulus situation is part of the organism's nature, then the whole relation (including its other term, the stimulus) seems to be brought *within* the organism, so that we cannot really understand the assertion that there *is* this relationship between distinct terms (this being the insuperable problem for Lewin's life-space).

This unworkable view that a thing can be in whole or in part made up of its relationships is the crux of hypostatization—of all doctrines of unseen forces or magical entities. This may be seen more clearly if we consider the organism's relations to its own responses. For the most part, a response really stands logically as a new property of the organism, i.e., as a change in its nature, and so of course it *is* a part of the organism. But we are concerned with the relationship of this new property to the state of the organism immediately *prior* to its appearance, and if we make that relationship a part of the preceding state of the organism, then we have the characteristic form of the fallacious doctrines we are discussing; that is, the organism produces that response simply because it is *in its nature* to act in that way—it has a propensity to do so. At one stroke we are absolved from seeking for those actual states of the animal which determined that, under some specific stimulation, it would produce that response, and from discovering precisely what the stimulus is, since it would seem merely an "inclining" cause at most, being in

fact a necessary but not sufficient condition.

This is the very fallacy MacCorquodale and Meehl are criticizing in their attack on the use of libido to explain features of behavior, but they mistakenly think it appears when libido is ascribed properties of its own, and that the way to avoid the fallacy is carefully to keep its nature devoid of any surplus meaning, i.e., surplus to "its" functions. The same prescription seems to be the one central to operationism, and though in the long run it goes astray, it is possible to see some force in it in one specific connection—namely, in the definition of dispositional concepts, with which in one place MacCorquodale and Meehl identify intervening variables as they conceive them. We must sympathize with at least the policy implied in Stevens' cry: "Only then shall we never think of energy or consciousness as a substance . . ." (15, p. 330).

Taking "solubility in water" as a dispositional concept, then following Bergmann's argument (3, p. 98), a proper (positivistic) definition of it would be to say, e.g., that "*x* is soluble in water" means "if *x* is put in water, it dissolves." The verification of the first sentence is held to be completely reducible to that of the second, which implies (and Bergmann makes this quite explicit) that the words in the first sentence mean precisely the same as the words in the second sentence. (Some positivists hold that this complete reducibility applies to *all* defined concepts, though that extreme view seems to have been modified by Carnap [5, p. 464, ff.]. The general proposition that all definitions are nominal comes down to saying that there is no such thing as coextension, which, as I suggested above, is tantamount to a rejection of determinism.) Now, it seems to me that the suggested definition is deficient in that it does not take into account ob-

jects which in fact would dissolve in water but which never are placed in water. (Modern symbolic logic would hold that the defining sentence could be converted to "either *x* is not put in water *or* *x* dissolves," and that if the first of these disjuncts is satisfied by *x* never being put in water, then the *whole sentence* is true of *x*—i.e., the conditions for *x* being soluble are satisfied, and the definition is held in this way to be adequate even for the negative instances. But if we admit that meaning of "either . . . or . . .," then, as Carnap points out [5, p. 440], the definition would include not only lumps of sugar that are not put in water, but also such things as a wooden match that is not put in water—anything, in fact, that does not meet that fate would "satisfy the conditions" for being soluble in water.)

The difficulty can be met, however, by amending the definition to read "*x* is of *such a nature* that if it is put in water, it dissolves," even though we do not know what that "nature," that common quality or character of soluble things, may be. Now, in principle, the verification of the presence of this character is not completely reducible to the observation that when *x* is put in water it dissolves, since any quality is theoretically observable, and its presence then could be established (if we knew what it was) by direct observation, without the necessity for observing its effects. But one might say roughly that the verification of solubility is thus reducible because (and this in my view is the point the positivists are really getting at here) it is not the character in question. To call a thing "soluble" is just to say that it has some unspecified character in virtue of which specified events produce a specified change in it, and that it retains that character even when those events do not materialize. In itself it merely poses the question, what this character is, but it is very fre-

quently used as if it were the answer to that question—as if it were that character itself—and that, in fact, is the typical way in which hypostatization occurs.

Ironically, however, by its very insistence that dispositional concepts are not “substances” and do not have “surplus meaning,” positivism sometimes leads to precisely the sort of mysticism it is trying to make impossible. That is, it is taken up wrongly (even by many positivists) as suggesting that there is *no* characteristic there in virtue of which the events in question take place, that the mysticism lies in going on to look for it, that the scientific procedure is to be content with “solubility,” and that we cannot fall into confusion as long as we rigidly exclude from our thinking any suggestion of quality or “substance” in the matter at all (failing to see that it is only from the notion of “solubility” itself that it must be excluded). But such a course of thought (which seems to be MacCorquodale and Meehl’s) makes the fallacy inevitable; if there is no relevant quality there, nothing which produces the dissolving but is describable in terms which *make no reference to* producing that effect, then that relationship (to dissolving) must be thought of as just “being in the nature of the thing.” Not only is such a notion not itself a solution, but while it is retained it specifically makes a solution impossible.

The reply might be made in defense of MacCorquodale and Meehl that they plainly do recognize the possibility of finding the qualitative processes mediating any response since they recognize “hypothetical constructs” as scientifically legitimate. But in my opinion the error remains; for them it is *not* that the hypothetical construct is found alongside of, and mediating, the intervening variable—not that the qualities which determine that given events pro-

duce given consequences still remain distinct from that relation between antecedents and consequences. Rather it is that “the existence propositions . . . automatically make the construct ‘hypothetical’ rather than ‘abstractive’” (11, p. 99)—i.e., the intervening variable *becomes* a hypothetical construct in the ascription *to it* of qualitative content. This is a further indication of what is made plain in their discussion of Skinner’s “emotion”: that the true intervening variable is thought of as a “state” *internal to* the organism, in some sense part of its constitution yet stripped of all qualitative content—and in that case it can only be the relativistic, mystical sort of notion that they are confusedly setting out to attack.

Although I said that there is one real set of facts indicated by some of MacCorquodale and Meehl’s uses of the term “intervening variable”—namely, the mathematical relations between types of event impinging on the organism and types of response produced by it—it does not seem to me that the term “intervening variable” is necessary or suitable for referring to it, partly because of other explicit meanings that have accrued to it, and partly because the words themselves inevitably suggest some state-like thing that *intervenes between* stimulus and response. A relation is not “between” its terms even in the most neutral way; one should say, rather, that they have, or stand in, that relation.

The mere giving of a name or symbol to those relations strengthens the ever-present temptation to slip into the “imaginary force” way of thinking, especially when the correlations are less than unity. In this case we cannot say that the stimulus in question produces the specific response (because there are some cases in which it does not), but if we remain convinced that there is some connection, then we are likely to say

that the stimulus results in a tendency to produce that response. But this tendency regularly finds its way inside the organism (because of a confused recognition that the state of the organism has something to do with it), and appears as a "demand" or "propensity" (or any of the multifarious species of "tendency") to make that response. (As "valence" it has found its way into the stimulus-object—i.e., the stimulus has "a tendency" to produce the response.)

Now, the use of "intervening variables" is sometimes defended by insisting that they *are* "imaginary" forces and are never intended as anything else. Their sole function is to help us grasp the observed facts and to organize our thoughts about them. But it is possible for our thoughts to be "organized" in a way that is mistaken or even meaningless. It may for the most part be true that the actual verbal forms are nothing but ways of expressing the observed connections, but the mere fact of offering "intervening variable" statements makes it vaguely seem that our knowledge is being extended, and that we are "getting to understand the facts better" in the sense of seeing how they are produced. The attribution of events to occult forces is rarely explicit because then it is so blatantly unscientific; it creeps in unacknowledged and gains its influence by default, as it were—by our failing to look for the *actual* causes.

To discover causal relations (which are always coextensive relations) we must take into account not only the nature of the forces acting on the thing in which the changes are produced, but also the properties, especially the fluctuating ones, of that thing itself (which in psychology will, for the most part, be, of course, the organism). When we discover which of its *actual* properties are involved in any given reaction, then

the need to fall back on imaginary forces whose sole function is to produce those effects (i.e., on "intervening variables") will have disappeared.

REFERENCES

1. ANDERSON, J. Realism and some of its critics. *Aust. J. Psychol. Phil.*, 1930, **8**, 113-134.
2. ANDERSON, J. The problem of causality. *Aust. J. Psychol. Phil.*, 1938, **16**, 127-142.
3. BERGMANN, G. The logic of psychological concepts. *Phil. Sci.*, 1950, **18**, 93-110.
4. BERGMANN, G. Theoretical psychology. *Annu. Rev. Psychol.*, 1953, **4**, 435-458.
5. CARNAP, R. Testability and meaning: I-III. *Phil. Sci.*, 1936, **3**, 419-471.
6. CARNAP, R. Testability and meaning: IV. *Phil. Sci.*, 1937, **4**, 1-40.
7. FEIGL, H. Operationism and scientific method. *Psychol. Rev.*, 1945, **52**, 250-259.
8. FREUD, S. Heredity and the etiology of the neuroses. In E. Jones (Ed.), *Collected Papers*. Vol. I. London: Hogarth, 1924. Pp. 138-154.
9. HULL, C. L. *Principles of behavior*. New York: Appleton-Century, 1943.
10. HULL, C. L. The problem of intervening variables in molar behavior theory. *Psychol. Rev.*, 1943, **50**, 273-291.
11. MACCORQUODALE, K., & MEEHL, P. E. On a distinction between hypothetical constructs and intervening variables. *Psychol. Rev.*, 1948, **55**, 95-107.
12. O'NEIL, W. M. Hypothetical terms and relations in psychological theorising. *Brit. J. Psychol.*, 1953, **44**, 211-220.
13. SKINNER, B. F. *Behavior of organisms*. New York: Appleton-Century, 1938.
14. SPENCE, K. W. The postulates and methods of 'behaviorism.' *Psychol. Rev.*, 1948, **55**, 67-78.
15. STEVENS, S. S. The operational basis of psychology. *Amer. J. Psychol.*, 1935, **47**, 323-330.
16. TINBERGEN, N. *The study of instinct*. London: Oxford Univer. Press, 1951.
17. TOLMAN, E. C. The determiners of behavior at a choice point. *Psychol. Rev.*, 1938, **45**, 1-41.
18. WATSON, J. B. Psychology as the behaviorist views it. *Psychol. Rev.*, 1913, **20**, 158-177.

(Received August 31, 1953)

RESPONSE FACTORS IN HUMAN LEARNING¹

GEORGE MANDLER

Yale University²

Theoretical treatments of human learning phenomena have been largely confined to an analysis of stimulus variables. In recent years, however, more attention is being paid to response factors, especially in the investigations of Underwood (28), Morgan and Underwood (20), and Osgood (22). This paper will present a theoretical framework, applicable to human learning and thinking problems, which will stress response factors. Attention will also be paid to the point, made by McGeoch (18) and others, that most "new" learning in human adults is at least partly a transfer phenomenon. The paper will be particularly concerned with the differentiation of stimulus conditions depending on the evocation of responses made by the organism, the relationship between overt and symbolic responses, and the transfer and overlearning of these responses. Thus, stimulus factors will be viewed as dependent upon the particular response repertory and previous experiences of the individual. This is in line with the position of Sperry (27) who has recently argued, from a neurological point of view, for a response approach to problems of perception and thinking.

The definitions and assumptions which represent the theoretical struc-

ture will be specified and some applications will be given, with particular reference to the effect of overlearning on transfer. The reader will note that some of the assumptions are deductive corollaries of others, or are related to other theoretical systems. For the present purposes they will be stated as assumptions.

DEFINITIONS

Stimulus. The term stimulus will be used as defined by Hull (15), essentially in terms of receptor input. Hull defined as "actual stimuli" those events which activate a receptor.

Overt response. This term will refer to any observable activity of the organism.

Symbolic response. A human organism will be considered to have made a symbolic response analogous to the overt response if he reports the perception of the overt response without performing the overt response.

Reinforcement. The term reinforcement will refer to an event in the environment which indicates to an individual the correct performance of a response. (No position is taken in this paper as to the specific mechanism of reinforcement. Presumably the above formulation is consistent with all current theories of reinforcement.)

ASSUMPTIONS

The Differentiating Response

1. A stimulus is differentiated from other stimuli when it evokes a response different from the response evoked by other stimuli. This differentiating response will be designated as R_s . The

¹This paper is part of a dissertation submitted to the Graduate School, Yale University, in partial fulfillment of the requirements for the degree of Doctor of Philosophy. The writer wishes to thank Drs. F. D. Sheffield, C. I. Hovland, and N. E. Miller for their suggestions and help in the preparation of this paper.

²Now at Department of Social Relations, Harvard University.

R_s can belong to any class of responses, i.e., it can be verbal, motor, or symbolic, depending on the original learning experiences of the individual.

2. When identical R_s responses have been frequently reinforced for two or more different stimuli, these stimuli, other things being equal, will be perceived as identical. Conversely, when different R_s responses have been reinforced to two or more different stimuli, these stimuli will be perceived as different. Identity or difference of stimuli, qua stimuli, depends solely on receptor stimulation, but identity or difference in the perception of these stimuli depends upon the differentiating responses. Thus identical stimuli cannot be perceived as different, but different stimuli can be perceived as identical. Another important implication is that two stimuli which differ in receptor stimulation, but do not evoke any differentiating responses, cannot be perceived as different.

3. Several different differential responses can be associated with any one stimulus and, other things being equal, they will differ only in terms of the probability of their evocation, which is a function of their reinforcement history, as in Hull's "habit family hierarchy" (14).

It will be noted that the present concept of differentiating responses is closely related to Dollard and Miller's (7) concept of cue-producing responses. Their description of attaching the same (or distinctive) cue-producing responses to distinctive (or similar) stimulus objects does not differ from the present treatment. The present statement does imply, however, that stimuli cannot be differentiated unless different responses are evoked.

The adult human organism responds to complex stimulus situations with a variety of previously learned responses which in turn become the stimuli in

the learning of new responses, i.e., they "mediate" the new associations. Thus, in a concept-formation experiment, the learning of a nonsense syllable response to all "green" stimuli is an example of mediated learning. On the other hand, a child's learning to differentiate the colors "red" and "green" (different stimuli) by differentially attaching the two verbal responses would be considered original learning.

Response Integration

1. Many responses performed by human organisms consist of aggregates of several subresponses which may be innate or acquired.

2. With successive repetitions of a response aggregate, the separate responses eventually become stimuli for each other such that any part of the response aggregate will tend to evoke the whole response aggregate. This process will be referred to as integration of the response.³

3. Integration is an increasing function of reinforced repetitions of the response aggregate.

4. The growth of integration is dependent upon the elimination of responses which prevent or delay reinforcement (as in "anticipatory errors").

5. The integration or association of two responses proceeds more rapidly than the association of a response with a stimulus. Thus, it is easier to learn a new response to a stimulus which already evokes a differentiating response than to a new unfamiliar stimulus.

6. The integration of a response aggregate proceeds more rapidly when the response units belong to the same effector modality. Differences in effector

³ Integration as used here does not refer to the simple chaining of responses, but rather to the simultaneous elicitation of an aggregate of responses. Responses that occur as overt chains very often are integrated, in the present sense, at the symbolic, perceptual level.

modality refer to differences in effector organs utilized in making the response. Thus it would be easier to integrate two verbal responses than a verbal and a motor response.

This conception of integration is closely related to Hollingworth's (11) concept of redintegration. Guthrie (9) has modified Hollingworth's concept in terms of parts of a response tending to condition each other.

Symbolic Responses

1. Any overt response which is perceived by a human organism evokes a symbolic response analogous to the overt response.

2. The symbolic response tends to be activated whenever the overt response is performed. Evocation of the symbolic response, however, tends to elicit the overt R_s only if motivation to perform the response is present.

3. Whenever a stimulus evokes two separate integrated responses, the two symbolic analogues may also be activated and associated so that, on future presentations of a stimulus which evokes only one of the responses, both symbolic responses will be activated.

4. Symbolic responses can be associated with other symbolic or nonsymbolic responses. In particular, they can be associated with verbal responses.

5. Previously learned verbal responses can have inhibiting effects which prevent occurrence of an overt response. When a particular overt response no longer leads to reinforcement, verbal statements as to its incorrectness can be attached by this experience to the symbolic analogue. On future occasions when the symbolic response occurs anticipatory to performing the overt response, the inhibiting verbal response can effectively forestall the error.

6. In a learning task, the symbolic analogue of a response aggregate will

differ from one trial to the next as long as irrelevant overt responses are still present and errors are still made. Thus, irrelevant responses which do not prevent, but are also not correlated with, reinforcement will from time to time be represented in the symbolic analogue. Necessary overt responses will continue to be represented and reinforced, while the irrelevant responses will drop out. In particular, the symbolic analogue is expected to be most distinctive and constant after many contiguous repetitions of the same response aggregate, i.e., after errors have been eliminated and overt performance has reached an asymptote.

The actual modality of the symbolic response is not relevant to the applicability of this concept. It appears advisable to leave stipulation of modality unspecified, which makes it possible to extend the concept of symbolic responses to both verbal and nonverbal behavior.

APPLICATIONS

In this section an attempt will be made to integrate, in terms of the present theoretical framework, some of the empirical results that have been obtained in studies of human learning. The studies quoted are intended to be representative of the empirical data in a particular area rather than exhaustive listings of these data.⁴

Learning of Differentiating Responses

It is to be expected that complex stimuli will evoke several differentiating responses previously learned to different parts of the stimulus pattern. At the same time, however, differentiating responses can be learned to the total stim-

⁴ In the present discussion, no statements have been made about the effects of differential motivation and related concepts. Such phenomena are presumably operative in addition to the effects discussed here.

ulus. Rossman and Goss (24) found that, while subjects used recently acquired verbal differentiating responses to distinguish stimuli in new paired-associates tasks, they looked for "identifying parts" of stimulus terms more frequently and found these more helpful. We would expect such a preference since highly overlearned responses to the identifying parts presumably have a higher degree of probability of evocation than recently learned nonsense syllables.

On the other hand, adult human behavior also provides many examples of the association of identical differentiating responses to different stimuli. Thus, the same word written in several different handwritings or in print will evoke the identical response and will be perceived as identical unless differentiating responses referring to the stimulus differences are actually evoked.

Differentiating responses will be associated not only with the experimentally controlled stimulus but also with other aspects of the total situation. The integrated response is reinforced in the context of the experimental situation, particularly in regard to the experimenter, the apparatus, and so forth. A recent study by Bilodeau and Schlosberg (3) showed more retroactive inhibition when the experimental room was the same for both tasks than when the interfering task was learned in a different room.

Prior training with attendant integration of a differentiating response is expected to shorten the learning process on future occasions when that response is used. Hovland and Kurtz (12) found that prior familiarization with nonsense syllables facilitated learning of lists using those nonsense syllables.

Concept Formation

Under this heading will be discussed only those aspects of so-called concept formation in which a subject is required

to learn a common response to a class of stimuli. The experimental situation has most frequently been exemplified by Hebb's experiments (10).

It is assumed that in the presentation of a number of stimuli, many differentiating responses will be evoked by each stimulus. The stimuli will have been predifferentiated to varying degrees, and different aspects of each stimulus will evoke different responses. In most cases these responses are verbal responses which also elicit their symbolic analogues. At the same time, the subject learns to make the new paired-associates response (nonsense syllables in Hebb's experiments). These new responses, which have never before been evoked in the presence of these particular stimuli, are now associated with the already learned symbolic responses. In the process of successive presentations, the symbolic response which corresponds to the "concept" will be associated more frequently than any other with the new response. Thus, if one of the prior differentiating responses is "face," then the new response (the nonsense word) will be associated with the symbolic analogue "face." In successive presentations of instances of this concept, this association will be reinforced so that, eventually, the evocation of the symbolic response "face" will also evoke the new *name* of the concept—the correct nonsense response. Figure 1 shows a schematic representation of this process; R_{sz} is the symbolic response common to the stimuli, and will be more frequently associated with the concept response than other differentiating responses. This is similar to Hull's (13) theoretical description of concept formation. The difference is that, in the present formulation, the concept response is (at least at first) associated with the respective differentiating responses rather than directly with the stimulus components.

A recent study by Baum (2) indicates

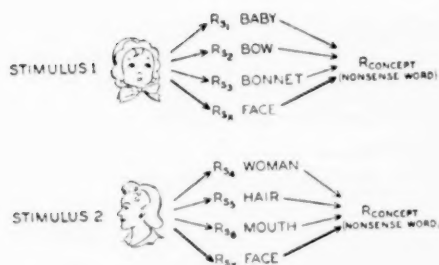


FIG. 1. Two instances of the same concept (Stimuli 1 and 2) eliciting one common ("face"), and several different, differentiating responses

that ease of concept attainment is a function of the discriminability of the stimuli. Her findings imply that degree of previous learning of the differentiating response determines ease of concept learning. In the extreme case, it would be expected that a stimulus which evokes no previously learned differentiating responses—the completely "unfamiliar" stimulus—could not be one of a class of "similar" stimuli. On the other hand, stimuli which have been maximally differentiated, i.e., with a high probability of evocation of a differentiating response, provide highly integrated responses, potent symbolic responses, and easy association with the new response.

Response Generalization

The general statement of this phenomenon usually implies that the learning of a response to a particular stimulus will facilitate the learning of similar responses to the same stimulus. This similarity can be specified in two dimensions. It can be either a similarity of overt parts of the two responses or a similarity of symbolic responses, as in the use of synonyms in Morgan and Underwood's experiment (20). An attempt will be made to show that response integration and symbolic responses are sufficient to explain the phenomena usually described as response generalization.

Similarity in terms of elements involves a communality of some of the parts of an integrated response. Substituting a new unit for one of the original units of the integrated response does not affect the integration of the units which are common to both aggregates. If it is assumed that the replaced unit can be dropped out fairly efficiently, this situation should produce faster learning than one in which all the units have to be integrated.

Similarity of symbolic responses, i.e., in the meaning realm, is comparable to the concept formation situation. The two synonymous responses both evoke a common symbolic response. As a concrete example, two of the synonyms used by Morgan and Underwood (20) were "dirty" and "unclean." To the extent that these two responses are associated with a common symbolic concept such as "filth," paired associate learning of the second response will be mediated by the common concept. Figure 2 diagrams this process. The common symbolic response, however, need not be as specific as the one used in the example, and it may be nonverbal. Morgan and Underwood also found that different degrees of synonymy are reflected in the degree of response generalization. The position taken here is that synonymy is a function of common symbolic representation of the two differentiating responses, which in turn would affect response generalization as found. A similar point of view, describing meaning as a function of commonly associated re-

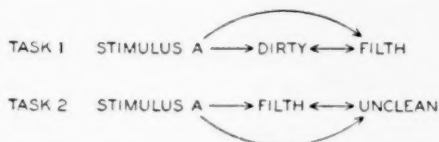


FIG. 2. The mediating and facilitating effect of a common symbolic concept in the response generalization of two synonymous responses

sponses, has been advanced by Noble (21).

Implicit Trial and Error

Dashiell (6), and others, have pointed out the importance of implicit trial and error in facilitating problem-solving behavior in human subjects. The facilitating mechanism is provided by the attachment of an implicit "yes" or "no" to the symbolic analogues of the various courses of action. A symbolic response sequence that leads to the anticipation of an unwanted consequence could produce one-trial inhibition in human subjects. The subject continues with such trial and errors—at the symbolic analogue level—until he hits upon the desired outcome; this sequence can then be translated into overt behavior.⁵

*Transfer of Training*⁶

Given an original learning situation, subsequent transfer situations can contain either new or the old stimuli, and either new or the old responses. "New" and "old" indicate whether or not the particular stimulus or response has been previously paired with any response or stimulus used in the experimental situation. Thus, a *new* stimulus is that stimulus which has not been previously used to elicit any of the responses used in the experiment, and a *new* response is a response or response aggregate which has not been previously reinforced in the experiment; *old* stimuli and responses are those which have been presented, or elicited and reinforced, in the original learning task.

⁵ The inhibiting "no" response would also produce other, primarily motivational, effects. Mechanisms such as fear and anxiety are probably additional factors in this inhibition phenomenon. It appears, however, that for the present purposes the more simplified conceptualization is adequate.

⁶ This topic will be treated more extensively than preceding applications, partly because relevant implications of the present theoretical approach have been explored experimentally (19) in a study to be published separately.

In line with the terminology used by Bruce (4), the four possible conditions of transfer are:

I. Making a new response to an old stimulus.

II. Making an old response to a new stimulus.

III. Making an old response to an old stimulus when these have not been previously paired.

IV. Making a new response to a new stimulus. This is considered chiefly as a condition in which the effect of previous training on new learning can be controlled; the kinds of "transfer" observed in this condition are not relevant to the present theoretical framework.

In order to schematize all four conditions, two *S-R* pairs are necessary in the training situation. This is forced by the third condition, where making an old response to an old stimulus requires more than one original pair. Thus, to make the other conditions comparable, two training pairs will be considered in each of the conditions. The schema is shown below, in which the original learning is S_1-R_1 and S_2-R_2 . New stimuli and responses are indicated by S_3, S_4 and R_3, R_4 respectively.

	Conditions			
	I	II	III	IV
Training	S_1-R_1	S_1-R_1	S_1-R_1	S_1-R_1
	S_2-R_2	S_2-R_2	S_2-R_2	S_2-R_2
Transfer	S_1-R_3	S_3-R_1	S_1-R_2	S_3-R_3
	S_2-R_4	S_4-R_2	S_2-R_1	S_4-R_4

For the present purpose, only familiar stimuli will be considered, i.e., stimuli with a high probability of evocation of a differentiating response. With less familiar stimuli, the predictions would be similar, but would involve the learning either of a response specific to the stimulus as its differentiating response or, with relatively unfamiliar stimuli, the adoption of the experimental re-

sponse as the differentiating response. In the present analysis, the experimental response is added to, and integrated with, the previously established differentiating responses to the stimulus.

In the original training task (i.e., in the acquisition of S_1-R_1 and S_2-R_2), response integration and generation of symbolic analogues proceed as a function of degree of training. Similarly, the probability of evocation of the response is increased. In the transfer situation, degree of training will produce differential facilitation or interference effects. Predictions for the three transfer conditions, schematized above, follow:

1. *Learning to make a new response to an old stimulus.* The relevant processes during original training are the integration of the response, the attendant formation of the symbolic analogue, and the increasing probability of evocation of the correct response. At high degrees of learning, the symbolic analogue will be distinctive enough so that on the first trial of the transfer task, i.e., the first time the subject makes the now incorrect old response, an "inhibitory" response will be attached to the symbolic response and will tend to decrease the probability of evocation of the now interfering response. Thus, at high degrees of overlearning on the original task, we would expect less negative transfer than at previous stages of training. Attaching the inhibiting response to both the overt response and its symbolic analogue on the first and succeeding transfer trials will tend to evoke the symbolic response and the inhibiting response before the overt response on subsequent presentations of the stimulus, and may thus avoid the overt response entirely. In this way the highly overlearned response has a sufficiently distinctive symbolic analogue so that the inhibiting response is readily associated, and the usual effects of associative in-

terference between the old and the new response are much less in evidence. The present analysis therefore predicts less interference with high degrees of learning.

A study by Underwood (29) confirms this prediction. Using adjectives (highly integrated responses) as stimuli and responses, he found a decrease in associative inhibition as the degree of prior learning was increased, and marked facilitation when the original task had been learned to a high degree. The present model would predict no facilitation when familiar stimuli are used, and Underwood's experiment is not crucial in this respect since he did not control for the effect of high degrees of learning on warm-up. Otherwise, the only effect to be expected is a decrease in associative interference.

In this transfer condition the latter effect would be expected to be less pronounced when nonintegrated responses are utilized, in which case more trials on the original task would be needed to show the decrease in negative transfer. Bruce's data (4), using nonsense syllables (rather than meaningful adjectives), do not show the decrease in interference at higher degrees of learning if his data are corrected for the positive transfer that he found in the new-stimulus new-response condition.

In the area of motor learning, Lewis, McAllister, and Adams (17) have shown that prior learning can produce both facilitation and interference. The task consisted of learning to manipulate controls which could be operated to superimpose two separate lights on a screen. On the transfer task, the controls were reversed so that, for example, a previously correct movement which had moved the light to the right now moved the light to the left. In this type of task, we would expect the similarity of the responses to produce some facilitation, while the reversal of the controls would produce interference. The data

suggest that, with increasing training on the original task, the number of correct responses on the reversal transfer task increases (response facilitation through transfer of a previously integrated response), but that specific errors also increase (interference due to previous reinforcement of the now incorrect response).

2. *Learning to make an old response to a new stimulus.* In this condition, the primary factor is assumed to be the integration of the response, i.e., the subject learns how to perform the response. Since the response has been learned not only in the presence of its original stimulus, but also is associated with all other constant aspects of the experimental situation, the probability of its evocation is high, and once evoked and reinforced in the transfer situation, the rate of learning of the new association will be partly a function of the integration of the response and its association with the context cues. It should be pointed out, however, that measurable transfer will reach a maximum before response integration does. For example, if the new association is learned on the first or second trial of the transfer task, increased integration of the response on the original task will show little measurable transfer effects. The prediction in this situation would be increasing positive transfer as a function of correct repetitions of the response on the original task.

The earliest relevant study is that of Bair (1), whose subjects learned to press specific colored typewriter keys to color stimuli. When a new list of stimuli was presented, with the responses remaining identical to the ones previously used, he obtained positive transfer. Bruce's data (4), again corrected for warm-up effects, show increasing positive transfer as degree of original learning is increased. Bunch and Winston (5) obtained clear positive transfer effects when a previously learned non-

sense syllable response had to be learned to a new stimulus.

The paucity of data in the literature relevant to this problem is compensated for by their unequivocal direction. However, one important implication from the present position has been given little experimental verification: if the response on the original task is already highly integrated (e.g., in the use of adjectives), then transfer effects should be minimal. In other words, amount of positive transfer in this situation is partly a function of degree of integration of the response at the beginning of the original task. The evidence that is available, however, is confirmatory (25).

3. *Learning to make an old response to an old stimulus when these have not been previously paired.* Three factors influence learning in the transfer task:

a. The integration of the response in the original task: facilitating effect.

b. The probability of evocation of the response, now incorrect, learned in the original task: interfering effect.

c. The symbolic "inhibition" of that incorrect response: counteracts the interfering effect of b.

At low degrees of learning, up to maximal strength of the original association, we would predict variable positive or negative transfer effects, depending on the relative strength of the facilitating and interfering effects *a* and *b*. When the symbolic analogue has become stable (at high degrees of over-learning), there would be an increasing tendency toward positive transfer as the interfering effect of *b* is counteracted by *c*.

One additional factor, however, is important in this particular condition. The differentiating responses evoked by the stimulus may be integrated to a greater or lesser degree with the response learned in the original task. In Conditions 1 and 2, this factor is presumably of minor importance. In Con-

dition 1, the old response is never reinforced in the transfer situation, and the integration would be additive to the general interference effect so that like-modality responses would lead to greater initial interference. In Condition 2, the effect might delay the increasing positive facilitation since part of the integrated response (the previously correct differentiating response) has to be eliminated. In Condition 3, however, these two effects would not only be additive, but the constant re-evocation of parts of this integration would interfere with the elicitation of either part alone. A recent study by Porter and Duncan (23) has shown greater negative transfer in Condition 3 than in Condition 1, when the stimulus and response elements were both verbal, which would favor integration of stimulus and response elements. In their discussion, the authors point to the possibility of the response re-evoking the stimulus and thus leading to greater interference.

If the differentiating response is verbal and the newly learned response (in the original task) motor (i.e., if the two responses belong to different effector modalities), we would predict that the integration of these two components would be minimal and would show less negative transfer than Condition 1. Siipola and Israel (26) have presented data bearing on this latter expectation. Subjects were pretrained to learn a series of responses on telegraphic keys. They were then presented with the original task in which these codes had to be associated with letters of the alphabet. In the transfer task, the same stimuli and responses were used, but their combinations were changed. The data, measuring transfer as a function of training on the original task, show slight initial negative transfer followed by a large positive transfer effect.

Kline's (16) subjects paired authors' names with book titles. He found that, with greater degrees of prior knowledge

of the *correct* authors' names, paired-associates learning was easier even when the correct response was to give wrong authors' names to the book titles. His evidence for decreased interference as a function of increasing familiarity is consistent with our prediction.

CONCLUSION

Further empirical verification of the above predictions should precede the application of this theoretical framework to more complex problems. The general emphasis on the model presented here has been on the importance of response factors in activities of the human organism prior to its introduction to an experimental situation. Phenomena such as stimulus discriminability are assumed to be a function of such previous experiences. In these terms, the differentiation of stimuli varies from individual to individual, and any general description of the discriminability of a stimulus only refers to the communality of experiences a group of individuals has had in learning differentiating responses to that stimulus. Thus, statements about stimuli, other than those referring to receptor stimulation, are limited to common social and learning experiences of subjects. In the past, studies such as Gibson's (8) have used stimuli and responses (with relatively homogeneous groups of subjects) which were most likely to result in similar differentiating learning experiences.

In reference to transfer effects, Guthrie's warning (9) that transfer is specific and not general has been taken account of. Particular attention has been paid to the fact that most human activities involve highly overlearned responses and response aggregates, and to the relevance of this phenomenon to transfer effects. If the predictions made concerning the differential effects of a subject's experiences with the stimuli and responses are borne out, then such long-accepted generalizations as Wylie's

(30), that "the transfer effect is positive when an old response can be transferred to a new stimulus, but negative when a new response is required to an old stimulus," need re-examination.

REFERENCES

1. BAIR, J. H. The practice curve: a study in the formation of habits. *Psychol. Monogr.*, 1902, **5**, No. 2 (Whole No. 19).
2. BAUM, MARIAN H. A study in concept attainment and verbal learning. Unpublished doctor's dissertation, Yale Univer., 1951.
3. BILODEAU, INA M., & SCHLOSBERG, H. Similarity in stimulating conditions as a variable in retroactive inhibition. *J. exp. Psychol.*, 1951, **41**, 199-204.
4. BRUCE, R. W. Conditions of transfer of training. *J. exp. Psychol.*, 1933, **16**, 343-361.
5. BUNCH, M. E., & WINSTON, M. M. The relationship between the character of the transfer and retroactive inhibition. *Amer. J. Psychol.*, 1936, **48**, 598-608.
6. DASHIELL, J. F. *Fundamentals of general psychology*. Boston: Houghton Mifflin, 1937.
7. DOLLARD, J., & MILLER, N. E. *Personality and psychotherapy*. New York: McGraw-Hill, 1950.
8. GIBSON, ELEANOR J. A systematic application of the concepts of generalization and differentiation to verbal learning. *Psychol. Rev.*, 1940, **47**, 196-229.
9. GUTHRIE, E. R. *The psychology of learning*. (Rev. Ed.) New York: Harper, 1952.
10. HEIDBREDE, EDNA. The attainment of concepts: I. Terminology and methodology. *J. gen. Psychol.*, 1946, **35**, 173-189.
11. HOLLINGWORTH, H. L. General laws of reintegration. *J. gen. Psychol.*, 1928, **1**, 79-90.
12. HOVLAND, C. I., & KURTZ, K. H. Experimental studies in rote-learning theory: X. Pre-learning syllable familiarization and the length-difficulty relationship. *J. exp. Psychol.*, 1952, **44**, 31-39.
13. HULL, C. L. Quantitative aspects of the evolution of concepts. *Psychol. Monogr.*, 1920, **28**, No. 1 (Whole No. 123).
14. HULL, C. L. The concept of the habit-family hierarchy and maze learning. *Psychol. Rev.*, 1934, **41**, 33-52.
15. HULL, C. L. *Principles of behavior*. New York: Appleton-Century-Crofts, 1943.
16. KLINE, L. W. An experimental study of associative inhibition. *J. exp. Psychol.*, 1921, **4**, 270-299.
17. LEWIS, D., McALLISTER, DOROTHY E., & ADAMS, J. A. Facilitation and interference in performance on the modified Mashburn apparatus: I. The effects of varying the amount of original learning. *J. exp. Psychol.*, 1951, **41**, 247-260.
18. McGEACH, J. A. *The psychology of human learning*. New York: Longmans, Green, 1942.
19. MANDLER, G. Transfer of training as a function of degree of response over-learning. *J. exp. Psychol.*, 1954, **47**, in press.
20. MORGAN, R. L., & UNDERWOOD, B. J. Proactive inhibition as a function of response similarity. *J. exp. Psychol.*, 1950, **40**, 592-604.
21. NOBLE, C. E. An analysis of meaning. *Psychol. Rev.*, 1952, **59**, 421-430.
22. OSGOOD, C. E. The similarity paradox in human learning: a resolution. *Psychol. Rev.*, 1949, **56**, 132-143.
23. PORTER, L. W., & DUNCAN, C. P. Negative transfer in verbal learning. *J. exp. Psychol.*, 1953, **46**, 61-64.
24. ROSSMAN, IRMA L., & GOSS, A. E. The acquired distinctiveness of cues: the role of discriminative verbal responses in facilitating the acquisition of discriminative motor responses. *J. exp. Psychol.*, 1951, **42**, 173-182.
25. SHEFFIELD, F. D. The role of meaningfulness of stimulus and response in verbal learning. Unpublished doctor's dissertation, Yale Univer., 1946.
26. SIIPOLA, ELSA M., & ISRAEL, H. E. Habit interference as dependent upon stage of training. *Amer. J. Psychol.*, 1933, **45**, 205-227.
27. SPERRY, R. W. Neurology and the mind-brain problem. *Amer. Scientist*, 1952, **40**, 291-312.
28. UNDERWOOD, B. J. *Experimental psychology*. New York: Appleton-Century-Crofts, 1949.
29. UNDERWOOD, B. J. Proactive inhibition as a function of time and degree of prior learning. *J. exp. Psychol.*, 1949, **39**, 24-34.
30. WYLLIE, H. H. Transfer of response in the white rat. *Behav. Monogr.*, 1909, **3**, No. 5.

(Received August 10, 1953)

KNOWLEDGE AND STIMULUS-RESPONSE PSYCHOLOGY¹

D. E. BERLYNE

University of Aberdeen, Scotland

A frequent source of uneasiness among psychologists is the increasing recklessness with which the stimulus-response type of theory, after years of servitude in animal laboratories, is being let loose among some of human psychology's most cherished preserves (11, 22, 37, 42, 48, 49). Many have long felt (46, 53) that even what lower animals do depends in some sense on what they "realize" about their environment, so that a psychology which does not have cognitions or perceptions as its basic concepts is poorly equipped for the study of infrahuman species. And many more can see, with a prophetic confidence paralleled only among the early opponents of experimental psychology, that S-R theory cannot progress very far with human behavior, because human beings do not just perform blind, automatic reflexes; they know what they are doing, and their actions are guided by what they know.

However, the only way to find out whether a theoretical approach is doomed to failure or is premature is to try it out. If we can find a way to analyze the role of knowledge and related phenomena in S-R terms, we may derive several important benefits. We shall be able both to take advantage of the rigor and precision of S-R language and to bring into view the relations between higher mental processes and fundamental principles of mammalian behavior. It has been abun-

dantly demonstrated that S-R terminology need in no way do violence to the flexibility and rationality of human activity. Accounts of reasoning have been offered (11, 26, 42) which do not appear to contradict radically either the facts adduced by those who have studied insightful problem solution or the tenets of S-R reinforcement theory. Discussion in similar terms of language processes, from which most of the psychological uniquenesses of human beings are generally held to spring, has been far from abortive (39, 49). And belated attempts to fit perception into the same framework have not encountered any immediate obstacle (3, 50, 52). The present paper discusses whether or not the human capacity for acquiring and using knowledge need escape the omnivorous maw of S-R behavior theory.

Since the concepts and principles which we shall apply to our topic originated in the investigation of very different problems, mostly in animal psychology, we shall have to rely on a procedure, which we shall call *concept extension*, that has often provoked misgiving. We shall propose that terms which were introduced to describe one class of phenomena—stimulus, response, drive, etc.—be extended to new classes of phenomena. Now it is frequently noted that human beings, and not least psychologists, are unduly prone to think that they have explained something when they have attached a new name to it. And cases of concept extension are apt to elicit the knowing look and the triumphant pounce from those who are commendably on the watch for such aberrations. But their well-meant vigi-

¹ This article arose out of work performed under the auspices of the Yale Communication Research Program, which is directed by Dr. C. I. Hovland and financed by a grant from the Rockefeller Foundation. It was written while the author was a member of the staff of Brooklyn College.

lance is here completely out of place, since concept extension is much more than mere labeling. The concepts which figure in systems such as Hull's behavior theory (29, 30) are defined by sets of relations with other variables. Applying them to new phenomena means therefore postulating that the same relations hold for these cases and thus laying down a set of hypotheses which can be followed until they prove inadequate. Concept extension is closely comparable to what happens in a court of law when it is ruled that taxis are hackney carriages or that gramophone records are a form of writing. This, far from being an idle verbal eccentricity, immediately applies a large body of traffic regulations or of law of libel to a large class of new instances, until further experience makes it desirable to pass new legislation.

"KNOWLEDGE" AS A CONCEPT IN BEHAVIOR THEORY

1. Ryle (47) has indicated that knowledge is a *dispositional concept*. An individual is said to possess it even when he is not in any way manifesting it, so that, like such terms as brittleness or electrical resistance, it expresses a probability that certain observable events ("truth-conditions" [10]) will occur, given certain additional conditions ("test-conditions" [10]). Dispositional concepts must, in a psychology which uses quantitative language, be represented by *intervening variables*, which are defined by describing the equations linking them to observable antecedent and consequent variables (29). At the present stage, our formulations cannot be too exact, since concepts only approach precise definition as a science progresses (10), but on the antecedent side, we can note that knowledge is a product of learning. Its strength depends on exposure to a stim-

ulus situation in the past, on the performance of responses in that situation, and on certain motivating and reinforcing conditions, which have been discussed elsewhere (4, 5). We assume that the possibility of innate knowledge, which at one time interested rationalist philosophers, can be discounted. On the consequent side, we can bear in mind Skinner's point that "to know is largely to be able to talk about" (49). It is true that testing ability to produce verbal behavior is one of the most convenient operations for measuring amount of knowledge, as in the traditional scholastic examination. But it is not the only one. There are times when it is dangerous to judge how much a person knows from how much he talks, and a dumb man may know more than a windbag. There appear to be, in fact, three principal effects of knowledge which can be a source of consequent variables: (a) performance of verbal responses, (b) production of new knowledge (as in reasoning) or evocation of other implicit responses (e.g., attitudes or thoughts), and (c) effects on overt behavior. As regards the last-named, all we can say at this stage is that knowledge causes the organism to behave in some respects as it would if the events or objects which are known were present. To use the terminology favored by cyberneticians and by Skinner, it causes present behavior to be "controlled" by absent or past events. With this point of departure, we can proceed to narrow down step by step the class of variables to which knowledge can be assigned.

2. The intervening variables that constitute knowledge are *habits*. The responses mediated by these habits may be overt, e.g., speaking or writing, or they may be implicit, e.g., thinking.

3. The responses mediated by these habits are *cue-producing responses* (11) and, if implicit, they are furthermore

pure stimulus acts (23). This means that they produce self-stimulation which can influence subsequent behavior of the same organism. Subjects who "know about" particular events or objects are said to "be conscious of" them. It is worth noting that, although consciousness no longer enjoys the consideration that was held its due in the days of introspective psychology, it is still found necessary to distinguish conscious from unconscious processes, because of the special properties of overt behavior dependent on conscious processes. Behaviorist writers (23, 35, 38) have attributed such behavior, regarded as largely a human prerogative, to the capacity to react to one's own reactions. Those (e.g., 11, 17) who have been inclined to follow Freud's assertion (14) that the unconscious is un verbalized have particularly stressed reaction to self-stimulation from verbal responses. The intervention of cue-producing responses has, moreover, been the key admitting insightful problem solution to the domain of S-R theory (11, 26, 42).

We have at our disposal two definitions of a "symbol" in behavioral terms, one offered by Morris (39) and one by Osgood (43). For the former, a symbol is a "*sign that is produced by its interpreter and acts as a substitute for some other sign with which it is synonymous,*" and "*if anything, A, is a preparatory-stimulus which in the absence of stimulus-objects initiating response-sequences of a certain behavior-family causes a disposition in some organism to respond under certain conditions by response-sequences of this behavior-family, then A is a sign*" (39, p. 10). This definition does not seem altogether satisfactory, as it would make signs of drive-stimuli, drive-producing stimuli and anything that induces any sort of set. Osgood's definition obviates this and other objections: "*a pattern of stimulation which is not the object is a sign*

of the object if it evokes in an organism a mediating reaction, this (a) being some fractional part of the total behavior elicited by the object and (b) producing distinctive self-stimulation that mediates responses which would not occur without the previous association of nonobject and object patterns of stimulation" (43, p. 204).

Osgood's scheme thus follows this pattern:



where $[S]$ is the sign, r_m the mediating reaction (a part of the response pattern made to the signified object), and R_X is the overt behavior evoked by r_m . It will be seen that r_m is the behavioral equivalent of the "concept" or "meaning" which a sign, in traditional accounts of symbolization, "conjures up" in the interpreter's "mind." Its introduction is necessary for one very good reason, apart from those put forward by Osgood. Most studies of symbols in animals and men have naturally begun with the single symbol as a unit. But in human higher mental processes, especially in knowledge, the units are actually complex combinations or sequences of elemental symbols, e.g., sentences (propositions) or complex perceptual responses. Since these combinations or complexes have effects that their constituents would not have alone, we have an instance of *patterning* as described by Humphrey (32) and analyzed further by Hull (28, 29). But this principle by itself will not suffice. The sentence "man bites dog" will have a different effect from "dog bites man" but a similar one to "human being sinks teeth into canine animal." Yet, one might expect the opposite to be true, since the first two sentences must be nearer together on any primary generalization continuum. Similarly, on the response side, the same remembered material is likely to be expressed in different words on different

occasions (2). So, different complexes of signs can come, through learning, to evoke the same "meaning" (which Osgood identifies by r_m), and the same meaning can come, through learning, to evoke different overt responses. We have therefore clear cases of "secondary stimulus generalization" (29) or "acquired equivalence of cues" (11) on the one hand and secondary response generalization, acquired equivalence of responses, or habit-family hierarchy (25, 30) on the other. These processes require a mediating cue-producing response (11, 25, 29).

When Osgood's r_m is evoked by a stimulus other than those coming from either the signified object or an external sign ("signal" [39]), whether such a stimulus be external or response-produced (e.g., by another r_m), we have what we shall hereafter call a symbol or symbolic response.

4. Knowledge mediates *believed* symbols. Morris (39) attempts a behavioral account of belief (which, he points out, can occur in differing degrees, as in judgments of probability) as the degree to which the organism is disposed to respond as if the signified object existed; Skinner (49) likewise defines belief in terms of strength of response. However, much of the response pattern conditioned to the significatum can occur even when the sign or symbol is disbelieved; it may very well evoke the same material in free association or directed thinking as it would if believed, and, as works of art and literature show, it may arouse similar emotional responses, though probably less intensely. It is rather in the *overt* behavior that we must look for a measure of belief, and it is principally this that is inhibited in doubt or disbelief.

Philosophers have, of course, long wrestled with the problem of distinguishing belief from knowledge. For some, knowledge is true belief, while others

insist that, in addition to being true, beliefs must be supported by adequate evidence to justify the name of knowledge. These questions, though important and no doubt capable of empirical formulation, need not concern us here. False beliefs affect behavior just as they would if they were true, at least until something arises to make the subject doubt them, and we can presume that similar motivation underlies the absorption of knowledge and error. We shall therefore not draw a distinction between knowledge and belief.

5. The believed symbols mediated by knowledge are *designative*. Morris (39) classifies symbols according to the distinguishing characteristics of the objects or events which they signify. If these characteristics are stimulus properties, the symbol is *designative*, if they consist of a preferential status with respect to the organism's needs, it is *appraisive*, and if they take the form of a tendency to evoke certain sorts of overt behavior, it is *prescriptive*. We can translate this valuable classification into Osgood's scheme by categorizing symbols according to which fractional component of the significatum's response pattern has come to form the r_m . If the r_m is composed mainly of emotional and drive-producing responses, it is an appraiser, whereas if it consists largely of fractional skeletal responses, it is a prescriptor. Designators will thus be those symbols which are built up of responses dependent on the stimulus properties of the significatum, and these may consist of perceptual responses (3) or verbal responses of the kind Skinner calls "facts" (49).

TRAINS OF THOUGHT

Many writers have described how the human being's responses depend jointly on the present stimulus situation and, with recent experience predominating, on a whole mass of relevant past experience which has left traces in his nervous sys-

tem. We have Herbart's apperceptive mass (21), Bartlett's schema (2), and the modern social and perceptual psychologist's frame of reference (20) as concepts referring to this phenomenon. Moreover, it has been pointed out that when these traces, which underlie knowledge, are reactivated, they give rise to long sequences of intraorganismic events—Bartlett's schema (2), Hull's sequences of pure stimulus acts (23), and Hebb's phase sequences (19).

There is no reason why we should not give to these sequences, which can bring about both the recall of old knowledge and, as in reasoning, the production of new knowledge, their everyday name, *trains of thought*. We can, by drawing on Hull (23, 29) and Bartlett (2), suggest the following six stages by which trains of thought may have developed out of simple response capacities in an animal as well equipped with ability to symbolize as the human being:

1. *Reaction*. Hull starts his account (23) by describing a series of events in the external world, which produce a parallel series of events or reactions in an organism. But, as he acknowledged in a footnote (p. 512), this account is deficient. The organism's reactions depend not only on external events but also on certain intervening variables representing conditions inside the organism; chief among these are the effects of previous learning (sH_R) and motives (D). This explains why different individuals not only perform different overt responses to the same external stimuli, but also derive different perceptions and later different knowledge from exposure to identical situations.

2. *Redintegration*. If S_1 , S_2 , etc. habitually occur in the same order or simultaneously, then foresight or expectancy can emerge. Both S_1 and s_1 (the proprioceptive stimulus resulting from R_1) can become conditioned to at least a fractional component of R_2 . These com-

ponents can include incipient skeletal responses (postural sets) and visceral responses such as fear (36, 41), but also, and these are what concern us most, perceptual and subvocal cue-producing responses.

If S_1 has come in this way to evoke, fractionally or subliminally, the perceptual response (\bar{r}_2) appropriate to S_2 (3), then some familiar phenomena from the psychology of perception follow from the principles of behavior theory.

- a. If the habitual S_2 follows or accompanies S_1 , then the principle of summation of reaction potentials (29, Corollary v) leads us to expect a lowering of the threshold for perceiving S_2 . There will be values of the reaction potentials $s_1E_{\bar{r}_2}$ and $s_2E_{\bar{r}_2}$ such that, although neither exceeds the reaction threshold (sL_R) separately, the behavioral sum of the two will. In any case, the sum of both will be greater than $s_2E_{\bar{r}_2}$ alone. Thus, this increase in perceptual reaction potential, which the writer has elsewhere identified with attention (3), explains the familiar fact that expected events are more likely to be perceived than others and are likely to be perceived more vividly. It is the phenomenon which Hebb calls the "central reinforcement of a sensory process" (19).

- b. If S_2 is an ambiguous stimulus, i.e., if it evokes two or more incompatible perceptual response tendencies of about equal strength, then the one reinforced by the expectancy conditioned to S_1 will prevail. This is one case of the influence of set on perception (9, 33, 52).

- c. If the habitual S_2 is for once replaced by a somewhat different stimulus, S_A , then several results might ensue:

- (i) *Illusion or dominance* (8, 9, 45). S_A will evoke \bar{r}_2 by stimulus generalization, response generalization, or both, although this will normally be less strong than \bar{r}_A , the "accurate" perception (i.e., the most frequent perceptual response to S_A). But \bar{r}_2 , when strengthened by

the expectancy aroused by S_1 , may well prevail, so that S_A will be wrongly perceived as if it were S_2 .

(ii) *Compromise* (8, 9). Both S_1 and S_A may evoke, by generalization, a perceptual response tendency corresponding to some stimulus occupying an intermediate position between S_A and S_2 on a continuum. In that case the strength of this compromise perception, contributed to by both, may exceed that of the expected perception (\bar{r}_2) and that of the accurate one (\bar{r}_A).

(iii) *Raised threshold*. If \bar{r}_2 is not strong enough to prevail over \bar{r}_A , it may interfere with it in such a way as to reduce its effective reaction potential ($s_A \bar{E}_{\bar{r}_A}$) (8, 45) and make its perception less probable.

d. If the habitual S_2 is absent, and the S_A which replaces it is so remote from it that no compromise or illusion is possible (i.e., because the generalized $s_A \bar{E}_{\bar{r}_2}$ is too weak), then two cases can arise:

(i) In conditions of poor visibility where S_A cannot be seen clearly (i.e., where $s_A \bar{E}_{\bar{r}_A}$ is weak), the \bar{r}_2 conditioned to S_1 may be supraliminal by itself. In that case we shall have a *hallucination* of S_2 . This happens when tachistoscopic figures are falsely completed (2) and when hallucinations are produced by suggestion (40) or conditioning (12).

(ii) In conditions of good visibility, where the discrepancy between S_2 and S_A cannot be overlooked, we shall have *conflict* between the incompatible perceptual responses, \bar{r}_A of peripheral origin and \bar{r}_2 of central origin. If we assume (7) that conflict is a drive condition (C_D), this explains the emotional effect that results from the clash between an expectancy and an external stimulus and plays a great part in Hebb's theory (18, 19).

3. *Symbolization*. Our consideration of stage 2 reveals the role played by previous knowledge in perception.² But

²It should be pointed out that, just as knowledge acquisition is best regarded as one

a step forward is achieved when two new conditions are fulfilled: (a) s_1 , the proprioceptive stimulus produced by R_1 , is sufficient without S_1 to evoke a fractional component of R_2 ; and (b) this fractional component can be supraliminal without support from S_2 . Then the fractional component of R_2 can become a symbol (r_m) for S_2 and represent it in its absence. We then have the possibility of a true train of thought, a sequence of internal responses (symbols) which can act in lieu of a remembered, anticipated, or imaginary series of external events. Each symbol is in its turn elicited by the response-produced cue of the previous symbol, so that a behavior chain (30) is formed, comparable to those of temporal maze habits (51, 54) or human rote memory (27, 31). The symbols (r_m) constituting such trains of thought may include perceptual responses (\bar{r}) or subvocal verbal responses (r_v).

4. *Ramification*. The next complications arise when S_1 participates in several habitual sequences of events at different times and thus can initiate several alternative associated responses. Similarly, each symbol in its turn may be able to lead the train of thought off in many alternative directions. But what determines precisely which response out of the many alternatives occurs? From Hull's account (30, p. 312) we can expect four factors to determine it jointly:

special sort of learning, namely that sort which enables a symbolic response to act in lieu of an absent stimulus, so the role of knowledge in perception does not exhaust the role of learning in perception. The cases we have been considering are those where the perception of a stimulus is supplemented or replaced by components of its perceptual responses which have been conditioned to cues habitually accompanying it. There appear to be many other cases where learning affects perception quite differently: a stimulus which could give rise to a number of alternative perceptions gives rise to one in particular because that one has been reinforced more than the others (34, 52).

a. *External stimuli* (S): In the case of autonomous trains of thought, only one of these is required in order to initiate the sequence. We shall therefore refer to such starting points as *initiating stimuli*.

b. *Response-produced stimuli* (s). These may be proprioceptive cues from muscular responses (s_p), or response-produced cues from perceptual responses (s) or verbal responses (s_v). They keep the train of thought going after the initiating stimulus has ceased.

c. *Drive-stimuli* (S_D). These continue throughout the sequence until the drive has been reduced and so become conditioned to every response in the chain. But the reinforcement-gradient principle implies that they will be most strongly conditioned to responses coming just before reinforcement.

d. *Fractional goal-stimuli* (s_G). These are internal cues produced by fractional anticipatory goal responses (r_G). They have the dual function of providing secondary reinforcement for earlier responses in the series (30, Corollary xv) and directing the series toward a goal (24).

The first two of these factors we shall call *cue-stimuli* and the last two we shall call *motivational stimuli*, noting that the two pairs have somewhat different roles. The cue stimuli provide the starting point for a train of thought and restrict the future course of the sequence to the relatively narrow range of responses to which they are conditioned. The motivational stimuli are conditioned to a much wider range of responses, since they must have coincided with an enormous variety of situations; they accordingly select from the repertoire made available by the cue-stimuli those items which are likely to contribute most effectively to the satisfaction of the motives aroused, and, in general, they serve to keep the train of thought on a path leading to the solution of the problem on hand. In addition, the drives with

which they are associated impel the chain of symbols to continue until the drives have been reduced or extinction has supervened.

The above conception has been derived from studies of maze learning in rats. It is therefore encouraging to note that other writers have been driven to recognize two corresponding sets of factors, as a result of direct attacks on higher mental processes in human beings.

Why only the correct association appears, whether it be a question of a single reaction, as in the controlled-association experiment, or of long successions of thoughts, as in directed thinking, was one of the principal interests of the Würzburg school. Their pursuit of the answer culminated in the theory put forward by Ach (1). It depends, he said, on the presence in consciousness of an "idea of the stimulus" (*Reizvorstellung*) and of an "idea of the aim" (*Zielvorstellung*). It is not hard to see in these two concepts the impact on the organism of cue-stimuli and motivational stimuli, respectively. They jointly produce a "determining tendency," which acts to steer the thought sequence toward the aim and to exclude irrelevant digressions.

Again, in Bartlett's theory of remembering, recall is the product of both the stimulus which elicits the remembering process (which "reminds" one) and what he calls an "attitude," which he describes as "very largely a matter of feeling or affect" (2). The latter ensures that the material which emerges is something pertinent to the present situation and not just a fortuitous association. Our cue-stimuli and motivational stimuli have thus obtruded themselves in yet another guise.

5. *Reorganization*. An important advance is accomplished when the symbols making up trains of thought are no longer tied to one chronological order but become capable of rearrangement.

This added flexibility makes possible "the assembly of behavior segments in novel combinations suitable for problem solution" (26; 30, ch. 10), and thinking can perform "the two different functions of preparation for reality (anticipation of what is probable) and substitution for reality (anticipation of what is desirable)" (13, p. 50).

The process of "short-circuiting" or "serial-segment elimination" presupposes, according to Hull (25), some persistent stimulus which acts during the whole of the sequence. Such a stimulus can become more strongly conditioned to later than to earlier items in the sequence, by virtue of the reinforcement gradient, and can thus serve to elicit anticipatorily those responses which immediately precede reinforcement, so that they crowd out irrelevant and unhelpful diversions. Internal events, and especially those we have termed motivational stimuli, serve this purpose.

Closely related conceptions appear in the writings of Hebb (19) and Bartlett (2). The former describes how the evocation of a familiar and long-established phase sequence comes in time to mean simply a review of its highlights, the less important connecting material gradually dropping out. Bartlett attaches an extreme importance to the ability of human organisms to "turn round on their own schemata," i.e., to "go directly to that portion of the organized setting of past responses which is most relevant to the needs of the moment" (2, p. 206). This ability obviates the necessity of reviewing a succession of trivial memories in order to reach the point of time which is important, as happens in some primitive forms of remembering. The factors responsible are "interest, appetite, etc." These are obviously motivational terms, and once again we can see an instance of motivational stimuli leading straight to those responses which are most "relevant" to them (i.e., most closely con-

tiguous with their cessation). Bartlett also describes the formation of specialized "schemata" (or organized system of retained material) pertaining to particular "appetites, instinctive tendencies, interests and ideals." Thus, once more we find attributed to motivational stimuli the power to tie together, and thus make readily available in close succession, those response tendencies which are most likely to subserve particular drives or purposes.

6. *Ratiocination.* The final refinement in the human being's application of knowledge is logical or, as Piaget (44) calls it, "operational" thinking. For this, the organism has to learn to perform only such symbol sequences as fulfil certain conditions ("rules of logic") which are necessary to ensure their stability and consistency. Some of these conditions are enumerated by Piaget, who outlines the stages by which a child gradually comes to achieve them. The reinforcement for this learning seems to come both from social reward and from the better adapted (more "intelligent") behavior that logical thought makes possible. Except when the restrictions of realistic reasoning are suspended—as in dreams and fantasy (16), wit (15), etc.—fortuitous, irrelevant, or illogical associations are inhibited. This is presumably because stimuli produced by such responses evoke some sort of acquired drive (e.g., Dollard and Miller's "learned drive to make . . . explanations and plans seem logical" [11, p. 120]).

If, as is hoped, this discussion violates neither the nature and importance of knowledge nor the findings of S-R learning theory, we can use the latter as a valuable source of hypotheses with which to attack many central problems in the higher mental processes. As an example, this account has given rise to a theory and to some experimental work on the much neglected topic of human curiosity, the motivation behind the acquisition of knowledge (4, 5, 6).

SUMMARY

An attempt is made to conceptualize knowledge in stimulus-response language. Knowledge, according to this analysis, consists of habits which mediate believed, designative symbols. It is suggested that symbol sequences or trains of thought are likely to have developed through six stages from the simplest response capacities to logical thought. Some of the phenomena that are familiar to investigators of thinking and perception are shown to be consonant with this account.

REFERENCES

1. ACH, N. *Über die Willenstätigkeit und das Denken*. Göttingen: Vandenhoeck & Ruprecht, 1905.
2. BARTLETT, F. C. *Remembering*. Cambridge: Cambridge Univer. Press, 1932.
3. BERLYNE, D. E. Attention, perception and behavior theory. *Psychol. Rev.*, 1951, **58**, 137-146.
4. BERLYNE, D. E. Some aspects of human curiosity. Unpublished Ph.D. thesis, Yale Univer., 1953.
5. BERLYNE, D. E. A theory of human curiosity. *Brit. J. Psychol.*, in press.
6. BERLYNE, D. E. An experimental study of human curiosity and its relation to incidental learning. *Brit. J. Psychol.*, in press.
7. BROWN, J. S., & FARBER, I. E. Emotions conceptualized as intervening variables—with suggestions toward a theory of frustration. *Psychol. Bull.*, 1951, **48**, 465-495.
8. BRUNER, J. S., & POSTMAN, L. On the perception of incongruity: a paradigm. *J. Pers.*, 1949, **18**, 206-223.
9. BRUNER, J. S., POSTMAN, L., & RODRIGUES, J. Expectation and the perception of color. *Amer. J. Psychol.*, 1951, **64**, 216-227.
10. CARNAP, R. Testability and meaning. *Phil. Sci.*, 1936, **3**, 420-471; 1937, **4**, 1-40.
11. DOLLARD, J., & MILLER, N. E. *Personality and psychotherapy*. New York: McGraw-Hill, 1950.
12. ELLSON, D. G. Hallucinations produced by sensory conditioning. *J. exp. Psychol.*, 1941, **28**, 1-20.
13. FENICHEL, O. *The psychoanalytic theory of neurosis*. New York: Norton, 1945.
14. FREUD, S. The unconscious. In *Collected papers*. Vol. IV. London: Hogarth, 1925. Pp. 98-136.
15. FREUD, S. *Wit and its relation to the unconscious*. In A. A. Brill (Ed.), *The basic writings of Sigmund Freud*. New York: Modern Library, 1938. Pp. 633-803.
16. FREUD, S. *A general introduction to psychoanalysis*. New York: Permapooks, 1953.
17. GUTHRIE, E. R. *The psychology of learning*. New York: Harper, 1935.
18. HEBB, D. O. On the nature of fear. *Psychol. Rev.*, 1946, **53**, 259-276.
19. HEBB, D. O. *The organization of behavior*. New York: Wiley, 1949.
20. HELSON, H. Adaptation-level as a basis for a quantitative theory of frames of reference. *Psychol. Rev.*, 1948, **55**, 297-313.
21. HERBART, J. F. *Psychologie als Wissenschaft, neu gegründet auf Erfahrung, Metaphysik und Mathematik*. Königsberg: Unzer, 1824-1825.
22. HOVLAND, C. I., JANIS, I. L., & KELLEY, H. H. *Communication and persuasion*. New Haven: Yale Univer. Press, 1953.
23. HULL, C. L. Knowledge and purpose as habit mechanisms. *Psychol. Rev.*, 1930, **37**, 511-525.
24. HULL, C. L. Goal attraction and directing ideas conceived as habit phenomena. *Psychol. Rev.*, 1931, **38**, 487-506.
25. HULL, C. L. The concept of the habit-family hierarchy and maze learning. *Psychol. Rev.*, 1934, **41**, 33-52; 134-152.
26. HULL, C. L. The mechanism of the assembly of behavior segments in novel combinations suitable for problem solution. *Psychol. Rev.*, 1935, **42**, 219-245.
27. HULL, C. L. The conflicting psychologies of learning—a way out. *Psychol. Rev.*, 1935, **42**, 491-516.
28. HULL, C. L. Words and their contexts as stimulus aggregates in action evocation. Unpublished memorandum, Yale Univer. Medical Library, 1941.
29. HULL, C. L. *Principles of behavior*. New York: D. Appleton-Century, 1943.
30. HULL, C. L. *A behavior system*. New Haven: Yale Univer. Press, 1952.
31. HULL, C. L., HOVLAND, C. I., ROSS, R. T., HALL, M., PERKINS, D. T., & FITCH, F. B. *Mathematico-deductive theory of rote learning*. New Haven: Yale Univer. Press, 1940.
32. HUMPHREY, G. *The nature of learning*. New York: Harcourt, Brace, 1933.

33. KILPATRICK, F. P. (Ed.) *Human behavior from the transactional point of view*. Hanover, N. H.: Institute for Associated Research, 1952.
34. KOHLER, I. Über Aufbau und Wandlungen der Wahrnehmungswelt. *Österr. Akad. d. Wiss., Phil.-hist. Klasse, Sitzungsber.* 227, 1 Abh., 1951.
35. LASHLEY, K. S. The behaviorist interpretation of consciousness. *Psychol. Rev.*, 1923, 30, 237-272; 329-383.
36. MILLER, N. E. Studies of fear as an acquirable drive: I. Fear as motivation and fear-reduction as reinforcement in the learning of new responses. *J. exp. Psychol.*, 1948, 38, 89-101.
37. MILLER, N. E., & DOLLARD, J. *Social learning and imitation*. New Haven: Yale Univer. Press, 1941.
38. MORRIS, C. W. Foundations of the theory of signs. *Int. Encyc. unif. Sci.*, 1938, 1, No. 2.
39. MORRIS, C. W. *Signs, language and behavior*. New York: Prentice-Hall, 1946.
40. MOWRER, O. H. Preparatory set (expectancy)—a determinant in motivation and learning. *Psychol. Rev.*, 1938, 45, 62-91.
41. MOWRER, O. H. A stimulus-response analysis of anxiety and its role as a reinforcing agent. *Psychol. Rev.*, 1939, 46, 553-565.
42. MOWRER, O. H. *Learning theory and personality dynamics*. New York: Ronald, 1950.
43. OSGOOD, C. E. The nature and measurement of meaning. *Psychol. Bull.*, 1952, 49, 197-237.
44. PIAGET, J. *La psychologie de l'intelligence*. Paris: Colin, 1947. (*The psychology of intelligence*. New York: Harcourt, Brace, 1950.)
45. POSTMAN, L., BRUNER, J. S., & WALK, R. D. The perception of error. *Brit. J. Psychol.*, 1951, 42, 1-10.
46. RITCHIE, B. F. The circumnavigation of cognition. *Psychol. Rev.*, 1953, 60, 216-221.
47. RYLE, G. *The concept of mind*. New York: Barnes & Noble, 1949.
48. SKINNER, B. F. *Science and human behavior*. New York: Macmillan, 1953.
49. SKINNER, B. F. *Verbal behavior*. (William James Lectures, Harvard University, 1947.) Cambridge: Harvard Univer. Press, in press.
50. SPENCE, K. W. Cognitive versus stimulus-response theories of learning. *Psychol. Rev.*, 1950, 57, 159-172.
51. SPRAGG, S. D. S. Anticipatory responses in serial learning by chimpanzee. *Comp. Psychol. Monogr.*, 1936, 13, No. 62.
52. TAYLOR, J. G. *The behavioural basis of perception*. In press.
53. TOLMAN, E. C. *Purposive behavior in animals and men*. New York: Century, 1932.
54. WOODBURY, C. B. Double, triple and quadruple repetition in the white rat. *J. comp. physiol. Psychol.*, 1950, 43, 490-502.

(Received August 13, 1953)

THE CONCEPT OF INTELLIGENCE AND THE PHILOSOPHY OF SCIENCE

CHARLES C. SPIKER AND BOYD R. McCANDLESS

Iowa Child Welfare Research Station

A careful application of the principles of the philosophy of science to controversial issues within an area of an empirical science has often proved clarifying. These methodological (logical) analyses have occasionally demonstrated that some of the questions which scientists considered appropriate for experimental attack could actually be resolved only after linguistic analysis. The major contribution of any such analysis is the reformulation of some of the traditional questions. The present paper attempts such an analysis of the psychological concept, "intelligence."

The paper is presented in two parts. The first contains a summary of the important points, relevant to this analysis, of the frame of reference within which the writers evaluate the methodological problems of their science. Writers of the philosophical school of "logical positivism," or "scientific empiricism," have written explicitly on the methodology of psychology, formulating principles that may be regarded as the fundamental principles of neo-behaviorism (2, 3, 5, 8). The second part deals with the application of these philosophical principles to problems associated with the investigation of human intelligence.

THE METHODOLOGICAL FRAME OF REFERENCE

The principles that scientists have followed in the formulation of their concepts have been made explicit by philosophers as a result of their analyses of the language of science, of which the language of the physical sciences is the prototype. The language of science is

a physicalistic language; that is, the referents of the descriptive terms occurring in scientific discourse are physical objects or events, their properties, and their relationships. There is, therefore, implicit in the philosophy of scientists a basic assumption regarding a "real world." The scientist assumes that there is a blueness "out there" when he has a sensation of blue. This "naive realism" of the scientist is not to be confused with any metaphysical viewpoints with reference to the nature of "reality." The scientist's position in this respect may be regarded as a convenient working assumption. It is simply another way of stating his belief that the data with which he deals have sufficient generality and significance to warrant further study.

Concepts that have been accepted in science and have proved useful for theoretical reasons, and for more pragmatic reasons as well, can be defined so that they are reducible to very simple terms, which have been designated by Carnap as *primitive predicates* (6). This class of terms is distinguished in part by the fact that they cannot be further reduced, in the sense that they cannot be given *linguistic* definitions; understanding of such terms can be obtained only through acquaintance with their referents. While philosophers have not troubled to delimit this class of terms categorically, its important characteristics may be given by a few examples. There are the property or quality terms such as "blue," "green," "bright," "hard," etc.; the relational terms such as "to the left of," "above," "between," "brighter than," etc.; and, of course, a

subclass of terms naming physical objects and events.

We may point out, parenthetically, that in scientific practice, concepts are not ordinarily reduced to (defined in terms of) such simple concepts. This would be laborious and, except for certain formal purposes, unprofitable. Words that may be reduced relatively easily to such a level are used without explicit definitions. Let us use the term "abstract" to refer to words whose definitional chains are long in the sense that numerous statements are required for defining them solely in terms of the primitive predicates. We may then describe scientific *practice* in this regard as that which utilizes explicit definitions only for the more abstract concepts; (even in these cases, the reduction process is carried down only so far as is necessary to avoid serious ambiguity). Such a statement, and rightly so, does not specify a crucial or necessary length of the definitional chain in order that the concept thereby defined be an abstract one.

If each acceptable term in a scientific language can be defined with reference to such terms as "blue," "above," "hard," etc., then the concepts in science refer in the last analysis to things that are *immediately observable* in a very simple sense of this italicized phrase. It is just this characteristic of scientific language which is intended when it is said that the language of science is a physicalistic language or that it has a physicalistic verification basis.

The formation of scientific concepts may be best understood through an exposition of the grammatical (logical) form of definitions in general, technically known as "definitions in use." Conventionally, one finds on the left side of an equation-like arrangement of two sentences a sentence in which the term to be defined occurs. This sentence ordinarily states one of the simplest things that can be said about the

term to be defined. For example, we may wish to define the concept "length." On the left side we may write the simple statement, "the length of this table is five feet." In the more important definitions, there is on the right-hand side of the definition a statement (or set of statements) that presents a relatively complex set of interrelationships among other terms, typically of the form: "If . . . then - - ." The two statements, the one on the right and the one on the left, are then connected by a symbol which carries the meaning, "means by verbal agreement the same as." If we fill in the right-hand side of the definition, the above statement about length means the same as: "If one takes a foot rule and repeatedly places it so that there is no gap and no overlapping of one placement and another, and if each placement is parallel to the edge of the table, then five such placements may be made between the edges perpendicular to the direction of the placements." The meaning of length is not explicitly carried, of course, unless the right-hand statement contains only terms which are already meaningful.

The groundwork has now been laid for an exposition of the phrase that has become so popular among psychologists—"operational definition." Bergmann (2) has pointed out that this term refers to nothing more complex than that science requires all terms occurring on the right-hand side of a definition to be, or to be reducible to, the primitive concepts we have already discussed. This requirement may be designated the *empiricist meaning criterion*, thereby avoiding some of the confusions which have become associated with "operationism." In order for a word to be meaningful by this criterion, it must be reducible, in the sense discussed above, to primitive predicates.

Obviously, this discussion describes an ideal procedure. One may look

vainly through some introductory physics textbooks for an *explicit* definition of the concept of "mass." What one ordinarily finds are several statements about mass, any one of which might, according to our discussion of meaning, be construed as a definition. This fact points up the need for considering the second methodological principle concerning scientific concepts. Analyses of scientifically acceptable concepts show that these concepts not only meet the empiricist's meaning criterion, but in addition are lawfully related to other meaningful concepts—such relationships being exemplified by statements of the form: "If A, then B," where A and B are both meaningful concepts. In general, the more relationships a given concept has to other concepts, the more fruitful or useful it is said to be. Thus, in physics, the concepts of time, force, energy, mass, distance, etc. are extremely useful since they enter in some form into all laws of mechanics.

Many discussions of operationism have been found objectionable by some scientists—particularly by some psychologists—because they have not emphasized this second aspect of scientific concepts. The scientist may insist that his term "means" more than just what is contained on the right-hand side of any definition of it. To anticipate later discussions, he may insist that intelligence means more than just an IQ from a given test: a high "amount" for a given individual means that this individual will probably do well in school, is probably good at arithmetic, is not likely to be found in an institution for the feeble-minded, probably has parents with high average school achievement, and the like. The present formulation does not rob the scientist of the richness of his "meaning." This *additional* meaning is carried by the statements of relationships between the unambiguously defined concept and the other concepts (i.e., school achievement,

institutionalization, level of parental education, etc.). Conveniently, Bergmann (2) distinguishes between meaning I (formal, operational meaning) and meaning II (significance, usefulness, fruitfulness). A concept that does not meet the first criterion cannot meet the second. A concept that meets only the first criterion will eventually be discarded as useless.

Since scientists are usually not so formal and explicit as are philosophers about such matters, one frequently finds in a scientific discipline useful concepts for which formal definitions have not been given. In some such cases, it is possible to formulate two or more equally correct and equally simple definitions. The question of which definition to select for a given purpose is therefore a matter of convenience. It is not consistent, *in a formal sense*, to speak of alternative definitions for a concept, since an unambiguous term can have only a single definition within the same context; but one may speak loosely of a number of concepts in science for which, in practice, several definitions are possible. This fact merely points out that it often happens in science that two or more grammatically different definitions may define concepts which are so highly interrelated that it is convenient to give each set of referents the same name. In other words, the relationships between each of these formally different concepts on the one hand, and other concepts on the other, are, within acceptable limits of error, identical. It makes little difference for most purposes which concept is used. A case in point is the concept of electric current, which may be quantitatively defined in terms of the deflection of a magnetic needle, the amount of heat generated, or the amount of silver deposited in a solution of nitrate of silver. When such clearly invariant relationships are found, it is often tempting (and, perhaps, of heuristic value) to

speak and think of the concept involved as if it referred to a "thing" ontologically independent of all the sets of operations, the description of any one of which could serve as the definition of the concept. It is usually implied in such discussions that the "thing" itself cannot be directly sensed, but that we "infer" its existence from the observable evidence (i.e., from the pattern of invariant relationships among the operationally defined concepts). Hence, it would be said, we may *measure* electricity, even though we cannot directly sense it, in much the same way that we might assemble evidence concerning the existence and size of a hidden room in a house by comparing external measurements of the building with measurements of the observable rooms in it. It should be apparent from what has been said previously that this is merely a manner of speaking, and like many metaphorical expressions, generates little confusion unless one begins to accept its literal meaning. In the latter event, scientifically sterile arguments arise as to what the "thing" would look like if we *could* directly sense it, or as to what the "correct" way is to measure (define) it.

It frequently happens in the development of a science that a word appearing in the everyday, common-sense language is taken into the language of that discipline and is given a new definition. In most such cases, the new meaning is in some sense similar to the meaning of the word in the ordinary language. The words "force" and "mass," for example, occurred in the English language before they were utilized in Newtonian physics. Most high school students of introductory physics learn to distinguish between the two meanings such words have, and little confusion seems to result. In the newer sciences, however, attempts are often made to convey factual information through the use of words from the ordinary language without explicit re-

definition of such concepts. In extreme cases, it appears that some scientists, particularly those in the social sciences, conceive of science as a technique for "measuring" the things to which many of the words in the ordinary language presumably refer. While it is not the writers' intention to depreciate the usefulness of common-sense knowledge, they wish to point out that if it had no limitations, scientific knowledge would not be necessary. Also, if the language of common sense were sufficiently precise, it would be unnecessary to study mathematics and logic. In many cases it appears that attempts to quantify (re-define) words from the natural language are uneconomical. Many such concepts refer in a vague way to highly complex sets of interrelations among distinguishable phenomena. It appears that the most economical way to study such patterns would be to define several concepts referring to these phenomena, with subsequent attempts to make explicit by empirical investigations the interrelationships holding among them. An all too frequent substitute for such a procedure consists of an attempt to "capture" all the phenomena and relationships in the definition of a single concept.

The complaint is not infrequently heard that if one subscribes to operationism, he places severe and perhaps crippling limitations upon the extent of the generalizations he can make. The argument proceeds along the following lines: Suppose a psychologist does a series of experiments on the learning of a task under certain conditions, using adult human subjects, and concomitantly defines a concept that he calls "habit₁." Operationally, the definition of this term includes references to the specific task, the conditions of learning, and the human subjects. Now, if he changes the task and the conditions, he must, according to the principles of operationism, define a new concept,

"habit₂." If he keeps the same task and conditions, but uses chimpanzees, he must again define a new concept, "habit₃." Obviously—the argument continues—such a procedure requires an inconvenient number of terms. Thus, operationism is too stringent and places too many restrictions upon scientific generalization. Since the business of science is the discovery of general laws, operationism defeats the purpose of science.

There are two distinct issues involved in the preceding argument. First, no one would argue that the subscripts to the above concepts do not have discriminable referents, and phenomena which *can* be reliably discriminated *may*, if one's purpose requires it, be given different names. Scientific practice may not typically be so formal as to apply subscripts to the terms, but it does differentiate among habits as studied in T mazes, in Skinner boxes, or in classical conditioning situations. Therefore, second, the question actually is whether a differentiation among such referents, either by name or by description, is a convenience or a hindrance. Concept analysis may be useful in pointing to the gaps in factual information where more careless terminological usage has obscured this lack. While it may point out logical differences among several concepts, it cannot indicate when there is sufficient empirical evidence to collapse these several concepts into a single one, or, more precisely, when it is possible and useful to define a more general concept which incorporates subsidiary concepts previously defined. Much of what is called theory in present-day psychology represents attempts to formulate more and more general concepts, whether they be called "habit," "drive," "aggression," "sign-gestalt-expectations," or what. In this last respect, scientists, without aid from the methodologist, are generally on guard against what Bergmann calls "that spurious compre-

hensiveness which is paid for by vagueness and triviality" (3, p. 438).

A similar objection to operationism probably arises from a failure to understand the formal (analytic) approach utilized by many writers in the exposition of this principle. The logician instructs us that a definition is arbitrary in the sense that it is the designation of a symbol (word) as a representation of an idea or complex set of ideas; which particular symbol is selected is of no formal importance; what is important is that the relationship between the word and its meaning be made clear and explicit. There is no empirical connection between a word and its referent. Objections to this formulation often take a form that suggests some type of word fixation or "concretism." It seems doubtful that such a mode of thought actually underlies many of these objections. What such people probably intend to emphasize—and logicians would be the first to agree—is that, in science, concepts are defined for some purpose. The scientist always wishes to define his concept in such a way that it will have a factual exemplification; that is, the referent of the term must exist in the same way that the referent of "chair" exists. Moreover, the scientist wants his concept to enter into statements of laws—in many cases, to enter only into certain laws. These two requirements depend upon factual matters for their realization. Thus, when the logician says that definitions are purely arbitrary, he speaks from a formal point of view and does not intend anything so nonsensical as that empirical considerations do not enter into the scientist's selection of a particular definition. It should be apparent that the answers to this objection, as well as to the one just previously stated, constitute restatements of the Meaning I—Meaning II distinction in slightly different guises.

The reader may note in this section of the paper an omission of any dis-

cussion of measurement and quantification. Since intelligence testing has been traditionally associated with such matters ("mental measurement"), this omission may be regarded by some as serious. The writers offer three reasons for their decision: First, the over-all logic of measurement, especially in psychology, has been clearly set forth by Bergmann and Spence (4). Second, the internal logic of test construction, together with its most widely accepted methods and techniques, has been comprehensively covered in such articles as that by Bechtoldt (1) and others. Finally, the writers consider this problem unessential to the understanding of the broader logic of the concept of intelligence, the primary concern of this paper. Misconceptions concerning the additivity of IQ points, the equality of units, the normal distribution of intelligence, etc. probably do not frequently occur among workers who are well grounded in the logic of statistics and measurement, and much of the confusion may be expected to disappear with improvement in such training.

THE ANALYSIS OF INTELLIGENCE

The term "intelligence" is one of a number of words that psychologists have taken from the natural language. Its common-sense meaning, like that of many similar concepts, is complex and indefinite. An unequivocal characterization of the common-sense notion is probably both impossible and unprofitable. Reflection on the common-sense meaning of intelligence, however, leads to the discovery of two important points: First, the meaning leads to logical contradiction since, on the one hand, an individual may be regarded as generally bright, and on the other, an individual may be considered intelligent with respect to one thing and unintelligent with respect to others. The second point is that the common-sense

meaning of intelligence always refers to behavioral consistency. There is the implication that the behavior of the individual is in some way trans-situational. Intelligence, in the common-sense usage, is not a momentary state of the individual, but transcends to some degree the specific situations in which the individual behaves.

In reading the nonexperimental ("theoretical") literature concerning intelligence, one must conclude that much time and energy have been devoted to attempts to capture and make explicit the several connotations of the natural language concept. Such attempts have probably stimulated much research. It is the writers' opinion, however, that numerous sterile controversies and confusions have arisen from an inadequate analysis of the goals and purposes of work on intelligence.

The organization of intelligence. There is one important assumption common to all the frames of reference in which intelligence tests have been constructed, from Binet to the present day. This is the assumption of trans-situational consistency of behavior. However, the different emphases of different test constructors have drawn attention to the inconsistencies of the original common-sense notion of intelligence. Some have argued that there is a general intelligence, that the trans-situational consistency in the level of behavior extends to all situations requiring "intellectual" problem solving. The term "intellectual" has actually been defined by the items selected for the tests rather than by attempts to circumscribe the "population" of intellectual behavior. But others, utilizing factor analysis as a tool, see no a priori limitations to the number of factors required to account for the variability of "intellectual behavior" (e.g., Thurstone [10]). For them, the empirical data determine the number of factors. Still another group of investigators has con-

sistently distinguished between "verbal" and "performance" intelligence, or between "abstract" and "concrete" intelligence.

It seems correct to state that no one, in any of these groups, has unambiguously circumscribed the population of "intellectual behavior" or has provided explicit sampling criteria for the selection of items for his tests. While this seriously limits the significance and objectivity of the frames of reference ("theories") in which the tests were said to be constructed, it does not detract in any way from any success in prediction that has been achieved by means of the tests; that is, the descriptions of the finished tests, and the accompanying instructions for administering and scoring them, constitute formally satisfactory definitions of the several concepts of intelligence, despite the lack of independent objective criteria for the initial selection of the items that constitute the tests.

The mathematical apparatus of factor analysis tends to obscure for some the fundamental logic of factor analytic investigations. The apparatus has been developed to handle simultaneously great quantities of interrelated data representing responses of individuals to test items. The completed analysis, if successful, indicates classes of test items that have elicited, within classes, similar responses from each individual in the sample, but on which similar responses have differed from individual to individual. The several empirically identified classes of items (stimuli) are then given names (e.g., "perceptual speed test," "number test," "test Y," etc.), and individuals receiving high scores on these classes are said to be high in "perceptual speed ability," in "number ability," etc. The prediction can be made that individuals from the appropriate population will tend to behave with intra-individual consistency on items within a class and will differ from each other in

the consistent mode of behavior within classes of items, and that relatively little consistency in behavior will be manifested from class to class. One of the presumed goals of this procedure is that tasks other than those used previously will yield to an objective analysis which will permit one to specify the combination of scores on the isolated factors that will be appropriate for successful performance on the task. Explicit rules for such analyses are not yet available. If such rules are ever specified, the utility of this approach will have been demonstrated.

Experimentation using factor analysis has attempted to study simultaneously groups of items toward which individuals behave with intraindividual consistency and with individual differences in the manner of responding to these classes of items. Except for the latter problem, the procedure does not differ in fundamental logic from the procedures that have been used to scale the psychological similarity of stimulus items. The meaning of the term "number test" or any other test can be given by stating the criteria for classifying the items into the test; this includes the entire factor analytic procedure. The meaning of the term "numerical ability" is given when the test is specified, the rules for administering it are given, and the scoring criteria stated. The term "factor" has often been used to refer to these different, though related, concepts. More "operational" definitions of psychological concepts could scarcely be given. It should be clear, however, that no "primary" factors, in the sense of physiologically or phenomenologically fundamental variables, can be said to have been isolated by the procedure utilized by the factor analyst any more than this could be said of other definitional procedures in psychology.

There is little sense to the question: "Which of these definitions of intelligence is correct (or most nearly cor-

rect)?" *Formally* correct definitions of all these concepts may be given. Which of the several concepts of intelligence proves to be the most useful, in the sense of entering into laws which lead ultimately to more accurate predictions of human behavior, remains to be seen. There is little use in speculating unduly on this point, considering our current state of ignorance concerning the variables associated with these concepts. Only empirical research can provide an unequivocal answer.

A similar analysis clarifies arguments concerning whether or not intelligence tests *need* to contain "nonintellective" items. We may recognize, first, that the occurrence of the terms "intellective" and "nonintellective" in everyday language does not guarantee that they refer to any features or phenomena that may be either consistently or usefully distinguished. If it is assumed for the moment that the terms are both useful and unambiguous, the proper question to ask is whether or not such items in a test will facilitate the achievement of the purpose for which the test was constructed. Test constructors are (understandably) rarely explicit about *all* the predictions they wish to make with their tests, and it is impossible to determine, *a priori*, whether or not any particular class of items will prove generally useful. Many of the controversial points concerning "the nature of intelligence" stem from an assumption that all investigators constructing or working with "intelligence tests" have a single common goal.

In this connection, Wechsler (10) asks if "the capacity for social adaptation" is not also a "sign of intelligence." He states that intelligence tests involve more than "mere learning ability or reasoning ability or even general intellectual ability." They also contain other "capacities which cannot be defined as either purely cognitive or intellective." He goes on to state that this

is desirable, and that such factors should be included with greater premeditation. One might well ask how one is to arrive at a sensible decision on this proposal until the goals of intelligence testing have been relatively clearly set forth. The issue, it would seem, is not one of a definition of an "absolute" intelligence that will be used generally; rather, it is necessary to state explicitly the criterion (or criteria) to be predicted, and then to discover the tasks that will predict it.

Heredity-environment. One of the most intense controversies in psychology in recent years was the heredity-environment issue. On the one side¹ was a group of individuals insisting that "intelligence" is something not directly influenced by the environment, i.e., not directly influenced by learning. On the other side, it was maintained that intelligence could be affected by learning experiences. This issue was closely related to the argument over the constancy of IQ, the insistence that IQ's obtained from certain tests (*viz.*, the Stanford-Binet) did or did not fluctuate markedly from time to time for a given individual. Reverberations of these controversies are still heard in current discussions of culture-free intelligence tests.

The salient points in this controversy were rarely, if ever, clearly and explicitly delineated. The polemical papers written on the subject indicate that much of the difficulty centered around careless use of terminology on both sides, and they suggest that a methodological analysis should prove clarifying. For example, the terms "environment" and "heredity" were never clearly de-

¹ The writers know of no reputable psychologist who could be said to belong unequivocally in one or the other of these mythical groups. Rather, the points at issue have been schematized in this way in order to represent more simply the pattern of the controversy.

defined, thus sharing the same ambiguity as "intelligence"—the concept they were intended to clarify. In the biological sciences, the term "heredity" is used precisely only in relation to the genotypically traced characteristics of the ancestors of the individual whose heredity is under discussion. Any attempts to define "intelligence" by referring to "heredity" would presuppose application of the procedures of the geneticist to the "intelligence" of the ancestry—and the circularity of this is apparent.

When one turns to research on the relationships between "heredity" and "environment" on the one hand and "intelligence" on the other, and construes these concepts operationally in terms of the research reports, one finds numerous definitions. A typical pattern of research was to provide an experimental group of children with specified experiences, to give pre- and posttraining intelligence tests, and then to compare the IQ gains with those of a control group not having the same intervening training. If greater gains occurred for the experimental group than for the control, it was held that the "environment" had influenced "intelligence." Few, if any, of these studies were devoid of serious experimental errors, the most damaging of which, in the writers' opinion, was the typical failure to assign subjects at random to the experimental and control groups. The foster home studies provide another pattern of research used by the "environmentalists," and were similarly limited by experimental errors.

The "hereditarians" had their own crucial experimental designs. If the IQ's for pairs of siblings reared separately correlated positively and significantly, it was the result of common heredity. If the IQ's for pairs of monozygotic twins correlated significantly higher than the IQ's for pairs of dizygotic twins, it was the result of more similar heredity for the former. Ques-

tions arise as to the importance of common uterine experiences, of the physical similarity of identical twins in leading to more similar environmental experiences, of the reliability in identifying identical twins except at birth, and so on. Jones (9) includes critical analyses of many papers in this area.

Much of the argument on the heredity-environment issue was not confined to such empirical questions as the foregoing paragraphs describe. Many workers in the area desired and expected a concept of intelligence which would provide a quantitative index that would not change with time for the individual except under the most unusual conditions, e.g., brain damage, psychosis, paralysis, etc. An intelligence test which suggested that intelligence fluctuated from day to day was therefore unsatisfactory; it was not a "real measure" of intelligence. The first empirical studies reporting systematic changes in IQ for groups were looked upon with considerable suspicion by many investigators. These studies and their supporters were answered with suggestions about uncontrolled variables that might have produced changes in IQ scores without affecting the fundamental intelligence. It now appears that this objection referred to the plausible possibility that IQ scores may be changed without materially affecting performances on tasks for which there was either a presumed or an experimentally established relationship with the IQ scores. The literature shows an interesting neglect of this possibility by those who insisted on the effectiveness of environmental factors in changing the level of intelligence. An obvious example of such a factor is coaching.

A terminological analysis helps to bring the conflicting conclusions into agreement. If intelligence is understood to refer to the performance on a given scale (Meaning I only), then without question, some environmental

influences (e.g., coaching, repetition of tests, etc.) can produce changes in intelligence. On the other hand, if intelligence is understood to refer to some complex set of interrelated behaviors (Meaning I and Meaning II), and if we have neither a complete list of the behaviors nor explicit statements of the relations holding among them, then we do not know and cannot determine whether or not learning experiences can produce changes in intelligence. As a matter of fact, if intelligence is understood in this sense, we can never know fully what intelligence "means," since subsequent investigations may uncover new relationships between the behavior and other concepts. One of the more important results of a methodological analysis of a scientific concept is the distinction made between the formal meaning of the concept and the empirical knowledge about the concept.

Analysis of the heredity-environment issue cannot be considered complete until mention has been made of the scientifically irrelevant values that have still further clouded the issues involved. The common-sense meaning of "intelligence" has a high value connotation for most of us, a characteristic it shares with many other psychological concepts (e.g., "rigidity," "neurosis," "prejudice," etc.). Intelligence tests have thus been evaluated by some, not only in terms of their predictive power, but also in terms of the "desirability" of the content. The evaluations seem to state: "Intelligence is 'good,' and if the test does not predict 'good' behavior, then it is not an intelligence test." This attitude often results in either a high evaluation of the IQ, per se, without adequate consideration for what can be predicted from it, or in bitter denunciation of test constructors who include questions in the test which handicap certain groups.

To ask whether it is good or bad for an individual to have high intelligence is about as scientifically relevant as to

ask whether it is good or bad to have an object weigh a lot. After scientists have defined their terms and have stated the interrelations among them, societies may decide whether or not a given term refers to something desirable. To reverse the procedure places on the scientist "pious" restrictions that are irrelevant to his purposes.

A survey of current literature on culture-free intelligence tests demonstrates this confusion of value and factual matters. For example, Eells *et al.* (7), with the most articulate of frames of reference, criticize the modern educational system and, therefore, the intelligence tests that predict success in it. They point out that middle-class teachers, with their particular middle-class version of what is the "best" and "true" culture, inflict their values upon school curricula, judgments of their pupils, and intelligence test items. Thus, they fail to develop the "full mental capacities" of their pupils, particularly of those pupils from lower classes. Present intelligence tests seek to predict behavior closely related to the school culture. They are, therefore, inadequate "to measure the general problem-solving activities of human beings." What is needed is an intelligence test that reflects or measures the "genetic mental equipment," "the general problem-solving activities," "the real talents," etc. Such an index would permit us to show that class differences in intelligence do not exist and thus help to prevent social class prejudice and untoward discrimination.

Without arguing for or against the educational goals of Eells and his co-workers, we make the following comments. Most psychologists would now agree that the predictive power of intelligence tests has been grossly overestimated, in both scope and accuracy, by many professional and nonprofessional people. But to criticize a test because it predicts one thing and not

another seems pointless. Whether or not a test can be constructed to predict important behavior, and yet not discriminate among social classes, is entirely a question of fact. Apparently Eells *et al.* (7) are attempting to construct such a test, and their attempts to make explicit the behavior they consider it important to predict should aid them. That part of their program concerned with a reformulation of educational goals can find no direct support from scientific knowledge since science cannot tell us what the "better life" is.

The validity of intelligence tests. Attempts to use technically the ambiguous term "validity" have generated much confusion in literature on intelligence. Consider the basic question, "Is this intelligence test valid?" One possible clear meaning of this vague question has to do with the usefulness of the test for predictive purposes. The answer to the question, by this interpretation, requires only a summary of the empirical research with the test. There is, of course, not much point in asking the question about a new test since little empirical knowledge will be available. If a new test is demonstrated to predict the scores on an older, well-established test, then the evaluation of the predictive power of the older test may be used for the new one. In this sense, the "validity" of a new test may be established relatively easily. Usually, however, the publication of a new test should be regarded as an invitation for other investigators to help to discover the predictive power of the test. If a given investigator judges that claims are made for the test that are not warranted by the empirical data, then it is his duty to register his objections. But a bland statement that the test is not valid contributes nothing but confusion and polemics to psychological knowledge. It amounts to nothing more than a forecast of future uselessness of the test.

The previous interpretation of the basic question has the virtue of permitting an eventual empirical answer. Another frequent interpretation is not so fortunate, having to do with whether or not the test is a *true* measure of intelligence. It presupposes a meaningful concept of *true intelligence*. It seems that such a question, unanalyzed, has led many workers to attempt to discover the "underlying nature of intelligence." It is rarely clear from their writings what is the "nature" of the "nature" they expect to find. It appears to have something to do either with the physiology or with the mental data of their subjects. The comments that follow are devoted to the issues that seem to be involved.

If one defines "intelligence" (or any other psychological concept) in terms of the individual's responses to items on a standardized test, one may still ask, "What are the physiological correlates of this type of behavior?" That every bit of behavior has physiological correlates is something of which psychologists are, as Bergmann puts it, "as certain of as we are of anything in science" (3, p. 442). Unfortunately, the more complex (molar) the behavior, the more likely it is that our present best attempts to specify which physiological variables underlie the behavior will be pure speculation and probably will be neither good psychology nor good physiology.

The problem is not greatly different in practice if one asks, "What are the mental correlates of this type of behavior?" No psychologist claims direct observation of his subject's mental data. If he is to do more than speculate, he must settle for observation of the subject's behavior (including verbal behavior) and the situations in which it occurs. He must assume that no mental states occur which are not *in some way* reflected in observable behavior.

The only important point that needs

to be made is that both the mental and the physiological correlates remain forever distinct from the behaviorally defined (psychological) concepts. Even if one finds an invariant relationship between a psychological and a physiological variable, they remain two things. One has found a law relating them. The failure to recognize this point has apparently led some writers to think of the physiological or mental variables as the "true" ones, which are only approximately "measured" by behavioral variables. What some psychologists seem to ask is whether or not the test reflects accurately the appropriate mental variables. The hopelessness of any immediate attempt to answer such a question is obvious. The most convincing answer one could give is the same answer one would give to the question, "How adequately does the test predict certain areas of behavior?"

To avoid misunderstanding, it should be made explicit that this formulation does not suggest that the study of the relationship between psychological and physiological variables is either an illegitimate or an unprofitable area for psychologists. Nor does it suggest that the study of subjects' verbal responses, under special instructional sets and conditions, as they relate to other situational or behavioral variables, is either a logical or factual error. The argument is merely that there are no a priori reasons why these variables are more fundamental ("real") than those at the behavioral level. This is a matter to be determined only by empirical trial and error.

SUMMARY

This paper is an attempt to examine some of the controversial issues in the field of intelligence by an application of some basic principles in the philosophy of science. A summary of the most

relevant of these principles was given, and the principles were then applied to such problems as the organization of intelligence, the heredity-environment issue, and the validity of intelligence tests. The aim of the analysis in each case was to separate terminological and other logical issues from the factual issues with which they have become confused. It was seen that there is little left that can be considered controversial, except in the sense that any question of fact may be a controversial point until adequate evidence is provided for its resolution. The confusions that arise as a result of trying to formulate single answers to multibarrelled questions can be eliminated.

REFERENCES

1. BECHTOLDT, H. P. Selection. In S. S. Stevens (Ed.), *Handbook of experimental psychology*. New York: Wiley, 1951. Pp. 1237-1266.
2. BERGMANN, G. The logic of psychological concepts. *Phil. Sci.*, 1951, **18**, 93-110.
3. BERGMANN, G. Theoretical psychology. *Annu. Rev. Psychol.*, 1953, **4**, 435-458.
4. BERGMANN, G., & SPENCE, K. W. The logic of psychophysical measurement. *Psychol. Rev.*, 1944, **51**, 1-24.
5. BRUNSWIK, E. The conceptual framework of psychology. Chicago: Univ. of Chicago Press, 1952. (*Int. Encycl. unified Sci.*, v. 1, no. 10.)
6. CARNAP, R. Testability and meaning. *Phil. Sci.*, 1936, **3**, 418-471; 1937, **4**, 1-40.
7. ELLS, K., DAVIS, A., HAVIGHURST, R., HERRICK, V. E., & TYLER, R. *Intelligence and cultural differences*. Chicago: Univ. of Chicago Press, 1951.
8. FEIGL, H. Operationism and scientific method. *Psychol. Rev.*, 1945, **52**, 250-259.
9. JONES, H. E. Environmental influences on mental development. In L. Carmichael (Ed.), *Manual of child psychology*. New York: Wiley, 1946. Pp. 582-632.
10. WECHSLER, D. *The measurement of adult intelligence*. Baltimore: Williams & Wilkins, 1944.

(Received November 4, 1953)

THE SKAGGS-ROBINSON HYPOTHESIS AS AN ARTIFACT OF RESPONSE DEFINITION¹

MALCOLM L. RITCHIE

University of Illinois

The literature on retroaction shows many variables involved in the determination of the experimental results. One of the most important of these is the similarity between original and interpolated tasks. Systematic investigations of similarity have been conducted since 1920. These experiments show that interference effects have either increased (e.g., 3, 7) or decreased (e.g., 4, 6) with increasing similarity between the tasks. Apparently, no experiment has reported reliable increasing and decreasing effects in the same experimental design.

There have been two major theoretical attempts (5, 6) to integrate the results of the similarity experiments to form a general similarity function. These general functions have assumed that the results of the many experiments are comparable even though they involve (a) many different definitions of similarity, and (b) more than one experimental design. It will be argued here that a confusion of experimental designs within given experiments has led to a procedural problem which is crucial to the interpretation of similarity results. This problem is a fundamental ambiguity concerning the definition of an acceptable response.

When the problem of response definition is recognized, two important consequences are apparent. First, a central issue in retroaction theory—the

similarity paradox—becomes a pseudo-problem. Second, one of the most frequently used measures of similarity—identical elements—is seen to be confounded.

THE SIMILARITY PARADOX

From early systematic studies of similarity, a generalization was drawn to the effect that interference increases as the similarity of the tasks increases. An extrapolation of this trend would show maximum interference when the tasks are made maximally similar. A paradox arose when it was noted that maximal similarity of successively learned tasks is the condition for ordinary learning; that is, continued learning with the same materials. Thus it appeared that the point of maximal similarity was at once the condition of (a) maximum facilitation, and (b) maximum interference. This is the similarity paradox to which Robinson (6) called attention and for which he proposed a resolution, which has come to be known as the Skaggs-Robinson hypothesis.

The Skaggs-Robinson resolution. Robinson began his analysis with the experimental data available to him, which showed increasing interference with increasing similarity. But he reasoned that there must be some point along this function at which the trend reverses, interference starts to decrease, and the function moves to high facilitation at the end point of maximum similarity. He also reasoned that there must be at the other end of the similarity scale a point at which no interference or facilitation is found as a function of complete dissimilarity of materials. The curve of this hypo-

¹This research was supported in part by the United States Air Force under Contract AF 33(038)-25726, monitored by the Air Force Personnel and Training Research Center. Permission is granted for reproduction, translation, publication, use and disposal in whole and in part by or for the United States Government.

thetical function goes from high facilitation (point *A* on Fig. 1) at maximum similarity to maximum interference at some intermediate point (point *B*), then to neutral at the point of complete dissimilarity (point *C*). Robinson had available evidence showing the slope from *B* to *C*. With a different experimental design he obtained results (6) giving the slope from *A* to *B*. However, many subsequent attempts have failed to produce the entire function within one experimental design and with one definition of similarity.

Osgood's transfer-retroaction theory. Osgood (5) pointed out the lack of specificity of Robinson's formulation. In a stimulus-response analysis of the experimental materials, he showed clearly that three basic types of experiments have been used: (a) stimuli constant and responses varied, (b) responses constant and stimuli varied, and (c) both stimuli and responses varied. The evidence, which Robinson tried to reconcile with ordinary learning (increasing interference with increasing similarity), Osgood holds to be the special case of simultaneous variation of both stimuli and responses. However, within this special case, Osgood's theory requires that the trend reverse itself in order to account for ordinary learning. The function thus obtained is very similar to that of Robinson.

This analysis of the experimental materials appears to leave the fundamental paradox unresolved. One fact that has

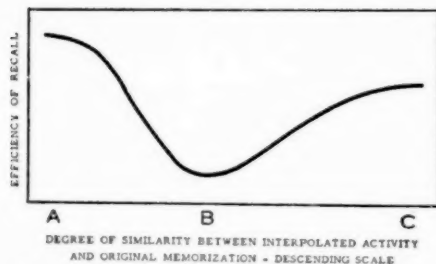


FIG. 1. The Skaggs-Robinson hypothesis. Hypothetical relation between similarity of original and interpolated material and amount of reproductive interference (6).

been overlooked is that there exists within these experiments a confusion of experimental designs. This confounding renders ambiguous the criteria by which the experimenter defines a correct response.

THE PROBLEM OF RESPONSE DEFINITION

In order for the results of different experiments to be combined in a general function, it must be assumed that the experimental procedures are comparable. When ordinary learning and interference data are compared, this assumption is violated. If we consider the criteria by which the experimenter defines a correct response, it can be shown clearly that two different procedures are involved. The procedure of the interference studies may be expressed as an ABA design, that of ordinary learning as an AAA design.

Response definition in the ABA design. The basic experimental design which has been used in the study of similarity effects is expressed in the ABA paradigm. Three learning series are involved: original learning (OL), interpolation of a different learning task (IL), and relearning (or recalling) the material of the original series (RL). The similarity between the original and the interpolated tasks is varied systematically. In the relearning series the subject is required to make the responses appropriate to the A series and not to make the responses of the B series. One of the stimulus-response paradigms for this design is as follows:

OL	IL	RL
S_1-R_1	S_1-R_2	S_1-R_1

The stimuli are constant for the three series and the responses are varied. It will be noted in this design that the subject must learn two different responses to functionally identical² stimuli. The

² The point has been made (5) that no two stimuli are ever absolutely identical. We may use the term "functional identity" to express our constant stimulus presentation.

performance on the relearning series involves a competition between the two responses. The subject must make R_1 and must not make R_2 in order for the trial to be recorded as correct. Regardless of how similar the two responses are, the subject is required to discriminate between them in order to satisfy the criteria of a correct response.

Now let us suppose that the experimenter increases the similarity of the required R_1 and R_2 toward the point at which they are functionally identical. As this point is approached, the subject has increasing difficulty in discriminating between them (the experimenter must retain a method of ready distinction), and the interference effects would thus be expected to increase. At functional identity the discrimination cannot be made. If the procedure of the interference design (ABA) is maintained, the interference effects between the two tasks are maximum—the subject cannot learn the two tasks at all.

Response definition in the AAA design. Ordinary learning proceeds on the basis of the successive presentation of stimuli and responses of functional identity. If we set up a paradigm for ordinary learning in the same manner as we have done for the ABA design above, our responses are distinguished only by the series in which they appear.

OL	IL	RL
S_1-R_1	S_1-R_2	S_1-R_1 or R_2

In this case R_1 and R_2 are functionally identical and are distinguished only by the series in which they appear. The subject is not required to discriminate between them in his performance. Either R_1 or R_2 is recorded by the experimenter as a correct response. This means that the subject does not have to discriminate between the response he made in OL and the response of IL. In this situation, facilitation is maximum at the same point as described above—

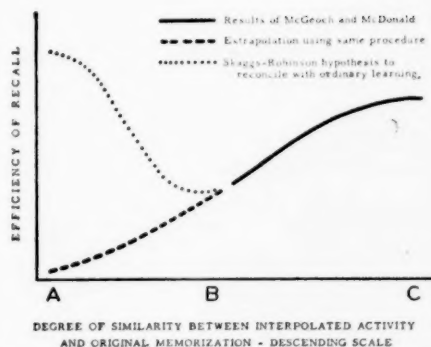


FIG. 2. Diagram showing results of McGeoch and McDonald (3), extrapolation involving continued use of ABA procedure, and the Skaggs-Robinson proposal to reconcile results with ordinary learning.

that of functional identity of both stimuli and responses. The difference between the two is the procedure by which the experimenter determines what is to be recorded as a correct response.

Robinson's analysis. The data that were available to Robinson were based upon the ABA procedure and showed increasing interference as similarity increased. The results of McGeoch and McDonald (3) have been used in Fig. 2 to represent this trend. Extrapolation of these results would give a prediction of maximum interference at maximum similarity. This extrapolation is based upon continued use of the ABA procedure and is shown by the dash line in Fig. 2. Robinson reasoned that there must be a reversal of this trend in order to account for ordinary learning (dotted line). Robinson's hypothesis requires a shift to the AAA procedure. The writer knows of no results in which this reversal occurred reliably when the same procedure was maintained.

METHODOLOGY AND THE DEFINITION OF SIMILARITY

From Fig. 2 it appears that the similarity function for the AAA procedure is considerably different from that for

the ABA procedure. It follows that any experimental design should involve only one of these for the results to be meaningful. The inclusion of both AAA and ABA in a given similarity experiment would confound the effects due to similarity with the effects due to the use of the two procedures. This is just the confounding that occurs when similarity is measured by the number of identical elements.

In the experiment of Robinson (6), for example, similarity was measured by the number of elements that were identical in the three series. The responses in his tasks have been schematized as follows:

OL	IL	RL	Elements in Common
a b c d	e f g h	a b c d	none
a b c d	a f g h	a b c d	one
a b c d	a b g h	a b c d	two
a b c d	a b c h	a b c d	three
a b c d	a b c d	a b c d	four

From the foregoing analysis it will be seen that wherever a response element appears in all three learning series, the AAA procedure is followed with respect to that element; that is, the subject is not required to discriminate between R_1 and R_2 . Wherever a different element is found in the interpolated task, the ABA procedure is followed; that is, only R_1 is acceptable as a correct response in relearning. Thus, in Robinson's experiment, increasing similarity was accompanied by an increase in the use of the facilitative AAA procedure. The less the interference procedure is used, the less would interference effects be expected in the results. This is precisely the result obtained by Robinson (6), Harden (1), Kennelly (2), and others using identical elements as a measure of similarity.

If varying the number of identical elements involves confounding similarity with the differential use of two procedures, then identical elements experiments do not give a legitimate similar-

ity function. Experimental techniques may possibly be devised to separate the similarity effects from the procedural effects. Until this is done, the results of identical elements experiments cannot be compared with those in which only one procedure is used.

SUMMARY

It has been maintained in this discussion that the similarity paradox in human learning was created by the analysis made by Robinson. Maximum similarity between OL and IL may be the condition for either maximum facilitation or maximum interference, depending upon the criteria established by the experimenter for defining a correct response in RL.

The evidence used by Robinson to support his general function is based upon an identical elements definition of similarity. Varying the number of identical elements was shown to confound similarity with the use of two response definitions.

REFERENCES

1. HARDEN, L. M. A quantitative study of the similarity factor in retroactive inhibition. *J. gen. Psychol.*, 1929, 2, 421-432.
2. KENNELLY, T. W. The role of similarity in retroactive inhibition. *Arch. Psychol.*, N. Y., 1941, 37, No. 260.
3. MCGEOCH, J. A., & McDONALD, W. T. Meaningful relation and retroactive inhibition. *Amer. J. Psychol.*, 1931, 43, 579-588.
4. OSGOOD, C. E. Meaningful similarity and interference in learning. *J. exp. Psychol.*, 1946, 36, 277-301.
5. OSGOOD, C. E. The similarity paradox in human learning: A resolution. *Psychol. Rev.*, 1949, 56, 132-143.
6. ROBINSON, E. S. The 'similarity' factor in retroaction. *Amer. J. Psychol.*, 1927, 39, 297-312.
7. SISSON, E. D. Retroactive inhibition: The influence of the degree of associative value of original and interpolated list. *J. exp. Psychol.*, 1938, 22, 573-580.

(Received October 19, 1953)

FIELD THEORY: II. SOME MATHEMATICAL APPLICATIONS TO COMPARATIVE PSYCHOLOGY

WILLARD E. CALDWELL¹

The George Washington University

Until vast amounts of research are carried out on animals other than the white rat and using drives other than hunger and thirst, and until all the quantitative interrelations are investigated, a field theory for comparative psychology will be relatively immature.

The concept of field has had many different applications in both physics and psychology. The attempt will be made here to point out some aspects of this concept as it is being used in this paper. Reference may be made to two previous articles developing this general frame of reference (1, 2). The organism as conceived here is a configuration of energy existing within a larger configuration of energy termed the environment. There is a constant interaction between the two fields. The organism is conceived of as following a process of differentiation with respect to the environmental field. This process reflects an attempt of the organism to achieve homeostasis with the environmental field. Examining a brightly lighted maze and a dark goal box may be viewed as a field of energy with a difference, and this difference may be measured by the intensity of light in the maze and in the goal box. The frame of reference used here is that the organism differentiates in the direction of less light, and in so doing it differentiates between the correct pathway and the incorrect pathway. In essence, then, a difference in the environmental field of energy produces a difference in the organism's field, as manifested in terms of time and errors. The field con-

cept is utilized here in a broader sense, inasmuch as the light is only one part of the total field of energy operating upon the organism; and we shall attempt to present a theoretical framework for isolating some of the major parts of the environmental field with respect to learning and perceptual problems.

In an earlier paper (3) the writer outlined a theory for, and some psychophysical techniques of, investigating maze learning when many different field-type drives are utilized. The concept of field drive is utilized somewhat synonymously with the term exteroceptive stimulus. It covers stimuli which originate outside the organism such as light, temperature, etc. This concept is used to cover types of motivation other than the internal biological drives such as hunger and thirst.

The first mathematical postulate in the earlier paper pertained to an experiment in which temperature and its reduction served as the motivating factor and as reinforcement. It dealt with the ratio between the increment of the difference between temperature in maze and goal box and the difference between temperature in maze and goal box necessary to produce a statistically significant difference in time and errors.

The object of this paper is to illustrate further how this psychophysical approach might be applied to maze learning, motivation, perceptual discrimination, and the Skinner-type design. The attempt will be made to illustrate it as an operational frame of reference applicable to exteroceptive types of stimulation and to many different types of animals.

¹ The author wishes to express his appreciation to Mary Lou Krehbiel for editorial assistance in the preparation of this paper.

TABLE 1

POSSIBLE COMBINATIONS OF EXTEROCEPTIVE STIMULI WHICH COULD THEORETICALLY BE TESTED PSYCHOPHYSICALLY IN MAZE, PERCEPTUAL DISCRIMINATION, AND SKINNER-TYPE DESIGNS

Apparatus (with modifications) and Locations of Measure of the Field	Fields of Energy	Formula for Varying both Measures of the Energy Field Concomitantly	Formula for Varying Stimulus in the Maze, in the Entrance Compartment, and before Pressing Bar	Formula for Varying Stimulus in the Goal Box or the Period after Pressing the Bar
Maze 1. Maze 2. Goal box	Temperature Light Revolutions Gaseous formaldehyde Angle of inclination Sucrose Amperage Sound Humidity	$\Delta(M-g) = K(M-g)$ $M = \text{Maze}$ $g = \text{Goal box}$	$\Delta(M-ge) = K(M-ge)$	$\Delta(Me-g) = K(Me-g)$
Perceptual Discrimi- nation 1. Entrance compartment 2. Goal box		$\Delta(E-g) = K(E-g)$ $E = \text{Entrance}$ compartment $g = \text{Goal box}$	$\Delta(E-ge) = K(E-ge)$	$\Delta(Ee-g) = K(Ee-g)$
Skinner Design 1. Before pressing bar 2. After pressing bar		$\Delta(B-A) = K(B-A)$ $B = \text{period before}$ pressing lever $A = \text{period after}$ pressing lever	$\Delta(B-Ae) =$ $K(B-Ae)$	$\Delta(Be-A) =$ $K(Be-A)$

SOME APPLICATIONS OF THE BASIC POSTULATE TO MAZE LEARNING

Table 1 presents an outline illustrating some of the permutations to which the writer's mathematical formulation would be applicable. The first major division of its applicability is maze learning. This postulate stated in abbreviated form is: *The increment of the difference between the stimulus in the maze and the stimulus in the goal box necessary to produce a just noticeable difference (j.n.d.) in time and errors is a constant fraction of the difference between the stimulus in the maze and the stimulus in the goal box.* The formula for this may be found in the third column of Table 1. The goal box and the maze both can vary. These formulae can be found in columns 4 and 5. Some preliminary work has been carried out to set up apparatus and procedures for testing some of the different types of stimulation utilized as motivation in maze learning.

Caldwell and Mosman (5) utilized temperature. Caldwell, Thaler, and Katz (10) performed a variation of the temperature-type experiment. Caldwell and Womack (12) utilized light avoidance. Caldwell and Sandler (9) utilized gaseous formaldehyde. Albino

mice were utilized in all five of these experiments.

Caldwell and Floyd (4) performed an experiment on albino mice placed in a maze which could turn a certain number of revolutions per minute and which would stop turning when the animals reached the goal box. This type of design lends itself to the possibility of varying the revolutions in the maze and in the goal box. Caldwell and Richmond (7) performed an experiment on hamsters wherein they utilized geotropism as the motivating factor. The maze had an angle of inclination of 21 degrees. The animals had to ascend the maze to reach the goal box, which had an angle of inclination of zero degrees. This experiment was also repeated by Caldwell and Ostrich (6) utilizing albino mice. This type of design also lends itself to the possibility of testing various combinations of differences represented by variations in the angle of inclination.

It is possible to use fish in field-drive experiments. An experiment was performed on goldfish which swam in a maze of high temperature to a goal box of low temperature (11). In this connection it is interesting to hypothesize an electrical field in a maze and either

its absence or reduction in a goal box. The maze itself might be positively charged and the goal box negatively charged. Various degrees of conduction might be applied in each.

The salmon has some photosensitive receptors deep in its skin (15). These are first covered by a layer of pigment, which subsequently disappears. As a result of this the fish reacts negatively to light. The hypothesis formulated here is that these receptors in the salmon could operate as motivating factors in the maze, and their reduction would serve as reinforcement. In testing this, various kinds of controls would have to be employed to separate the skin receptors from those of the eye.

Guttman (14) conducted some interesting experiments on rats in which he used bar-pressing responses reinforced with sucrose solutions of various combinations. This suggests the question of how much increment in sucrose is necessary to get a difference in rate of responding in the maze situation. The problem might be stated as the increment of the difference between the sucrose concentration fed before starting the maze and that fed in the goal box necessary to produce a j.n.d. in time and errors representing a constant fraction of the difference between the concentration fed before starting the maze and concentration fed in the goal box.

The stimulation of sound and humidity might be applied to this type of design and the various combinations of differences tested.

SOME APPLICATIONS TO PERCEPTION PROBLEMS

In Table 1 the same approach might be applicable to problems of perception. Flynn and Jerome's study (13) with rats might be applicable here. Light avoidance has been utilized with pigeons on perception problems (8). The light avoidance was employed for motivation in training pigeons to dis-

criminate geometrical figures. The apparatus consisted of a box which was brightly lighted and painted with aluminum paint. The goal box was relatively dark, and the goal-box doors were a circle or a triangle, depending upon the problem. If a pigeon entered the circle, for instance, the light would be turned off and the bird would remain in the dark goal box for five minutes. Error and time curves were established for these pigeons.

Our intention is to emphasize the possible applicability not only of light avoidance but of other types of field drives to problems of perceptual discrimination, and also to urge the psychophysical treatment of data derived from such types of experiments. The part of the apparatus where the animal is placed to make the discrimination is referred to as the entrance compartment, and the darkened area is designated as the goal box. The stimulation can vary in three ways similar to those suggested for maze learning.

For purposes of clarification, this problem might be stated as follows: *The increment of the difference between the light in the entrance compartment and the light in the goal box necessary to produce a j.n.d. in perceptual differentiation (in time and errors and correct choices) is a constant fraction of the difference between the light of the entrance compartment and that of the goal box.*

SOME APPLICATIONS TO THE SKINNER-TYPE DESIGN

Another problem is that of using field drives such as temperature, gaseous formaldehyde, light avoidance, etc. in the type of experimental design outlined by Skinner (15). Skinner's method may possibly be more sensitive to psychophysical measures than those of the standard maze-learning phenomena. Also, it is important theoretically to know how the results obtained from

utilizing the Skinner-type design and field drives compare with results derived from the use of maze-learning designs and field drives.

In order to test the following drives, the Skinner-type design must be modified, but the essential elements in the design should be retained—mainly, the instrumental one where a stimulus is introduced and the organism presses a bar to stop the stimulus. Records should be kept on the relation between the variation in the intensity of the stimulus and the variation in bar pressing. Guttman investigated bar-pressing responses in the rat where sucrose was utilized as reinforcement. He says:

Evidence is presented that rate of responding in the Skinner box with rats is a semilogarithmic function of the concentration of sucrose used as reinforcement. Extrapolation of the fitted rate-concentration function yields an estimated reinforcement threshold in the region of the sucrose-preference threshold and the human sucrose limen. Extension of this experimental technique to other reinforcing agents may yield a systematic pattern among reinforcement thresholds (14, pp. 360-361).

The application of this approach to the Skinner design might be further clarified by the following: *The increment of the difference between the intensity of light in the Skinner bar-pressing apparatus before the bar is pressed and the intensity of light there after the bar is pressed necessary to produce a j.n.d. in the time and frequency of bar pressings is a constant fraction of the difference between the intensity in the Skinner bar-pressing apparatus before the bar is pressed and the intensity there after it is pressed.*

SOME APPLICATIONS TO MOTIVATION

A fourth type of apparatus to which this general theoretical approach might be applied is that which attempts to measure motivation. There are many types of apparatus which are utilized for measuring activity levels. The revolving drum is one that might be applied here to these various exteroceptive stimuli. The difference in the stimulus field

could be measured with light, as an example, by measuring the light intensity in the animal's cage and then measuring it in the revolving drum. The j.n.d.'s would be in terms of activity level measured in terms of the number of revolutions of the revolving drum. This also raises the question of measuring the animal's frame of reference before placing it in the other types of designs mentioned in this paper. This may be one of the advantages of utilizing exteroceptive stimulation rather than hunger or thirst in animal experiments.

DISCUSSION

The foregoing programmatic outline of research is presented in broad outline form. The j.n.d. is actually a statistically significant difference in time, errors, and correct choices. Different parts of these curves obtained should be compared statistically. The quantitative results expected might appear only in experiments with certain animals. Perhaps only certain drives will be of use from a psychophysical point of view, possibly in connection with only a few types of animals, but ascertaining such facts requires that many animals be utilized in testing each variation of the hypothesis.

It may be that many of the experimental designs given here are too variable. The Skinner-type design was referred to for use in testing some hypotheses, but perhaps additional apparatus, more sensitive and of a new type, should be devised. Certainly the apparatus suggested here should be modified for the different species and for measuring such stimuli as light in comparison with the more conventional types of motivating stimuli such as hunger and thirst. Control groups are necessary where there is no difference between the entrance box and the goal box (or its equivalent) with respect to the particular drive being tested.

Details for investigating these abbreviated hypotheses must be worked out for each experiment. Reference should

be made to the writer's previous paper (3) on the application of psychophysics to learning and reinforcement.

SUMMARY AND IMPLICATIONS

Historically, the communication of ideas in a form in which the experimentalist can investigate them in the laboratory has been one of the principal functions of psychological theories. There are dangers and limitations in miniature quantitative theories, but there is also value in operationally defining problems so they may yield data that can invalidate or substantiate the basic assumptions underlying a theory.

This paper has urged that, from the psychophysical point of view, a tremendous amount of research is needed in the field of comparative psychology before we can begin to construct theories that possess any degree of maturity, either qualitatively or quantitatively. It also has attempted to present research problems that might be quantitatively tested in connection with the comparative aspects of motivation and reinforcement.

The implications are that further integration of the field drives with perceptual-type experiments, maze learning, motivation, and utilization of more sensitive techniques similar to Skinner's should be attempted in such a way that they may: (a) be checked psychophysically, (b) be checked with many different species, (c) have their functions checked against results obtained from the more conventional maze-learning experiments, and (d) yield results which might aid in giving us a more unified field theory for comparative psychology.

REFERENCES

1. CALDWELL, W. E. Adaptive conditioning: a unified theory proposed for conditioning. *J. genet. Psychol.*, 1951, 78, 3-37.
2. CALDWELL, W. E. The theory of adaptive differentiation. *J. Psychol.*, 1951, 31, 105-119.
3. CALDWELL, W. E. The mathematical formulation of a unified field theory. *Psychol. Rev.*, 1953, 60, 64-72.
4. CALDWELL, W. E., & FLOYD, J. P. The performance of albino mice in the maze situation with stimulation of the vestibular sense as motivation and its relative absence as reinforcement. *J. genet. Psychol.*, in press.
5. CALDWELL, W. E., & MOSMAN, K. F. The role of temperature change as reinforcement. *J. Psychol.*, 1951, 32, 231-239.
6. CALDWELL, W. E., & OSTRICH, R. The performance of albino mice in the maze situation utilizing gravitation and the vestibular sense as motivation. *J. genet. Psychol.*, in press.
7. CALDWELL, W. E., & RICHMOND, R. G. The performance of hamsters in the maze situation utilizing gravitation and the vestibular sense as motivation. *J. genet. Psychol.*, in press.
8. CALDWELL, W. E., & RICHMOND, R. G. The utilization of light avoidance as motivation in the investigation of perceptual differentiation in the pigeon. *J. genet. Psychol.*, in press.
9. CALDWELL, W. E., & SANDLER, H. M. The role of gaseous formaldehyde as reinforcement in maze learning in albino mice. *J. Psychol.*, 1952, 33, 47-56.
10. CALDWELL, W. E., THALER, W. D., & KATZ, J. J. The utilization of temperature change as motivation and reinforcement in the maze performance of albino mice. *J. genet. Psychol.*, in press.
11. CALDWELL, W. E., & TIEDEMANN, J. G. The performance of goldfish in the maze situation with the utilization of temperature as motivation and its reduction as reinforcement. *J. genet. Psychol.*, in press.
12. CALDWELL, W. E., & WOMACK, H. The performance of albino mice in the maze situation with the utilization of light as motivation and its relative absence as reinforcement. *J. Psychol.*, 1953, 35, 353-360.
13. FLYNN, J. P., & JEROME, E. A. Learning in an automatic multiple-choice box with light as incentive. *J. comp. physiol. Psychol.*, 1952, 45, 336-340.
14. GUTTMAN, N. Theories of reinforcement and the reinforcement threshold. *Amer. Psychologist*, 1953, 8, 360-361. (Abstract)
15. ROULE, L. *Fishes: their journeys and migrations*. New York: Norton, 1933.
16. SKINNER, B. F. *The behavior of organisms*. New York: Appleton-Century, 1938.

(Received October 10, 1953)

CRITICAL COMMENT ON "LEARNING AND THE PRINCIPLE OF INVERSE PROBABILITY"

ROBERT P. ABELSON

Yale University

An impression that the theorems of inverse probability are of widespread applicability to learning theory is created by David Bakan in his recent paper "Learning and the Principle of Inverse Probability" (1). His treatment is very simple and very ingenious and, if one were not to give the matter too much thought, his conclusions would seem quite sound and powerful. Many statements about learning rates, extinction rates, trial-and-error learning, insightful learning, etc. are made in Bakan's paper. The writer has no quarrel with these statements as such; however, it is felt that Bakan's statements do not follow from his premises. The subsequent discussion will attempt to show that inverse probability is not especially cogent to learning theory and that its use in that context is a misrepresentation either of inverse probability or of learning theory.

Bakan defines three entities: g , h , and x . For simplicity's sake, our discussion will include only g and x . However, it should be understood that h , which represents the ability level and prior experience of the organism, is assumed to be known in all of the subsequent definitions. The symbol g is defined as a certain state of the organism, presumably a state in which the organism is capable of responding in a particular way—the organism, when in the state g , might be said to be "knowledgeable." The symbol x is defined as a particular proposition of knowledge, a hypothesis about the environment (e.g., "If I press the bar, I will get a pellet of food").

Bakan's basic equation requires the following definitions:

$P(g)$ is the probability that the organism is in the condition g .

$P(g/x)$ is the probability that the organism is in the condition g after x is verified or reinforced.

$P(x/g)$ is the probability that x will occur if the organism is in the condition g .

$P(x/\bar{g})$ is the probability that x will occur if the organism is *not* in the condition g .

R is the ratio $\frac{P(x/g)}{P(x/\bar{g})}$.

Then Bakan writes (legitimately so):

$$P(g/x) = \frac{R P(g)}{R P(g) + [1 - P(g)]}$$

From this equation, Bakan derives all his results.

This equation is said to involve inverse probability because it attempts to infer causes from the observation of effects. The equation contains expressions for the probabilities of occurrences which are never observable [$P(g/x)$ and $P(g)$]. These probabilities ordinarily cannot be verified by counting the relative frequencies of favorable occurrences. Indeed, the philosophical dispute from which Bakan takes great pains to dissociate himself is concerned with the question of whether there exists any sense at all in which $P(g/x)$ can be considered a probability.¹ The negative position, tersely stated, is that "either the organism is in the condition g , or it isn't. A probability statement

¹ Carnap (3) is engaged in an extensive logical analysis of probabilistic statements. His work may provide a resolution of the long-standing philosophical dilemma.

is inappropriate." According to this position, many classes of phenomena would be excluded from the realm of discourse of probability. For instance, one would never speak of the probability of the truth of Weber's Law, or of the mass-energy relation $E = mc^2$, or of Freud's theory of unconscious motivation. These laws or theories are (provided the context of their application is defined) either true or they are false. Nor is there any connection between "approximately true" and "moderately probable" insofar as theories or laws are concerned.

From the point of view of the statistician, the only appropriate use of probability with these classes of phenomena is to fall back on the statement: "If I claim this theory to be true, the probability that my claim will be proven correct is such-and-so." Then if many claims are made, one can calculate the expected number of correct claims. This procedure can lead to a maximization policy with respect to claims about unobservable causes—a kind of static "game theory" approach to the collection of knowledge on the nature of the universe. "Maximum likelihood estimation," together with the method of "confidence intervals" in statistical theory (4, pp. 507–513), is such an approach.

Regardless of the philosophical merits of the argument against the use of inverse probability, its cogency in the case of Bakan's derivations is apparent. The argument is particularly devastating when it is realized that the only reasonable interpretation of Bakan's learning curves is that they are functions describing the experimenter's "game." If we take seriously Bakan's formulation, which allows only two conditions for the organism, g and "not- g ," then we will find that the learning curve for an individual organism is *not* a gradually rising curve. If we know which condition the organism is in at every trial,

we will find that the learning curve approximates a step function. When the organism is in the state \bar{g} , it will respond at a low probability level of success, $P(x/\bar{g})$, and will keep responding at this low level until suddenly it attains the state g . At this point, the organism will abruptly start responding with a high probability level of success $P(x/g)$ and will forever after maintain this high level. All learning would then be "insightful." Now, conceivably, Bakan's "learning curve" can be viewed as an average of an infinite number of such step functions, each with a different time of cross-over from \bar{g} to g . However, such a "learning curve" would have no meaning when applied to a single organism. To apply Bakan's gradually rising curve to a single organism is simply to admit our ignorance of the *actual* state of the organism at any given time.

Bakan exposes himself all the more to this criticism by using $R = \frac{P(x/g)}{P(x/\bar{g})}$ as a parameter of the learning curve, implying that in a given case one might measure both $P(x/g)$ and $P(x/\bar{g})$ in order to be able to specify the exact form of the learning curve. However, such a measurement can only be made if there is some criterion by which we can determine whether an organism is in g or in \bar{g} so that the relative probabilities of the occurrence of x in these two circumstances can be determined experimentally. But if such a criterion existed, then the sensible thing to do would be to apply it to the organism while it was learning to find out when it was in g and when it was in \bar{g} , and thus solve the problem at once. R , then, is a parameter with the following properties:

- Either 1. It can never be measured in practice
- Or 2. Its measurement destroys the theoretical grounds on which it is based.

Bakan's learning theory is not a theory of the learning process in a given organism; it is a theory of the process of analyzing the learning process of an organism. As such, it is typified by the situation in which a scientist analyzes the advancing state of his own knowledge. It does apply to "the method of science as a way of learning," as Bakan claims, but it applies only in this situation and not to classical learning theory. A rat is certainly not capable of analyzing his own learning process in this complicated way. To extend the results to classical learning theory would constitute a gross misunderstanding.

The current mathematical models for the learning process (Mosteller and Bush [2]; Estes [5]) *put the variable ignorance into the organism itself*, instead of into the experimenter. They assume not two states of the organism, g and \bar{g} , but a continuum of possible states p , where p is the probability that the organism will make the correct response. The hypothetical learning curve is given by p as a function of the number of trials. In this type of model, *the*

organism itself gradually becomes more and more certain of the correct response. In Bakan's model, the *experimenter* gradually becomes more and more certain that the organism is cognizant of the correct response. The former seems to be much the more appropriate model for the typical learning situation. That is not to say, of course, that Bakan's results cannot prove to be of value in the limited context of scientific method as the experimenter's way of learning.

REFERENCES

1. BAKAN, D. Learning and the principle of inverse probability. *Psychol. Rev.*, 1953, **60**, 360-370.
2. BUSH, R. R., & MOSTELLER, F. A stochastic model with applications to learning. *Ann. Math. Statist.*, 1953, **24**, 559-585.
3. CARNAP, R. *Logical foundations of probability*. Chicago: Univer. of Chicago Press, 1950.
4. CRAMER, H. *Mathematical methods of statistics*. Princeton: Princeton Univer. Press, 1946.
5. ESTES, W. K. Toward a statistical theory of learning. *Psychol. Rev.*, 1950, **57**, 94-107.

(Received January 8, 1954)

A METABOLIC INTERPRETATION OF INDIVIDUAL DIFFERENCES IN FIGURAL AFTEREFFECTS

MICHAEL WERTHEIMER AND NANCY WERTHEIMER

Wesleyan University

A decade ago, Köhler and Wallach (6) studied the effects of prolonged figural stimulation on certain subsequent perceptual test patterns, a phenomenon they called figural aftereffects. To account for these effects they postulated a change in the polarizability of that part of the brain upon which the previous contour had been projected, a process they termed satiation.

It has been repeatedly observed (2; 3, pp. 202, 207; 4, p. 202; 5, pp. 300, 316; 8; etc.) that there are individual differences in the size of such figural aftereffects. If one tentatively accepts Köhler's physiological model (3, 4, 5, 6), then such individual differences must reflect differences in the ease with which a modification in cortical conductivity can be brought about. Such a view implies that figural aftereffects could be used to measure generalized cortical modifiability in an individual. Assuming that this modifiability is characteristic of the entire brain, and not specific to a given area, this leads to the prediction that the size of figural aftereffects in different modalities should be correlated, i.e., a small kinesthetic figural aftereffect indicates low cortical modifiability, which in turn would predict a small visual figural aftereffect. Specifically, (a) visual and kinesthetic figural aftereffects measured on a large number of people should show a positive correlation (as also suggested in 4, pp. 196-197) and (b) intraindividual changes in visual and kinesthetic figural aftereffects should pursue a parallel course, assuming that an individual's cortical modifiability will change through time.

The satiation theory, as well as the newer statistical theory (7), involves physicochemical alterations in cortical tissue. These could be interpreted as implying metabolic changes. Thus one could argue that a relatively large figural aftereffect reflects high physicochemical modifiability and hence conceivably a high "metabolic efficiency."¹

Although this latter term is not sufficiently defined, it is adequate to yield some further predictions: (c) Size of figural aftereffect should correlate with physiological indicants of metabolic efficiency such as basal metabolic rate, thyroid activity, and such indices of circulatory efficiency as capillary structure. (d) It should similarly correlate with behavioral indicants of neural efficiency such as reaction time and ease of simple sensory-motor learning. (e) An experimentally induced alteration in metabolism should be reflected in a concomitant change in the size of figural aftereffects. (f) Schizophrenics, as a concrete example of a group of subjects with generally low metabolic efficiency (1, 8), should exhibit smaller figural aftereffects than normal subjects.

All the above predictions have been subjected to at least a preliminary empirical test, and all have been essentially confirmed, with the exception of the one concerning simple sensory-mo-

¹Originally we used the term "metabolic rate," but experimental evidence has shown that this concept seems to be inadequate. In our data, figural aftereffects are maximal in the normal range of metabolic functioning and fall off on either side. This result has tentatively led us to the term "metabolic efficiency."

tor learning, where the evidence was ambiguous.

Although these predictions are the only ones tested thus far, the present interpretation could yield many more, especially in classes *c* and *d* above, e.g., predictions concerning hormonal balance, stress effects, problem solving, and perceptual rigidity. Further, the vagueness of the present formulation has the virtue of making it compatible with any theory of figural aftereffects in which metabolic changes in neural tissue can reasonably be assumed, be the theory in terms of homogeneous conductors (6) or neural elements (7).

REFERENCES

1. HOSKINS, R. G. *The biology of schizophrenia*. New York: Norton, 1946.
2. KLEIN, G. S., & KRECH, D. Cortical conductivity in the brain-injured. *J. Pers.*, 1952, **21**, 118-148.
3. KÖHLER, W. Relational determination in perception. In L. A. Jeffress (Ed.), *Cerebral mechanisms in behavior: the Hixon symposium*. New York: Wiley, 1951. Pp. 200-230.
4. KÖHLER, W., & DINNENSTEIN, D. Figural aftereffects in kinesthesia. In *Miscellanea Psychologica Albert Michotte*. Louvain, 1947. Pp. 196-220.
5. KÖHLER, W., HELD, R., & O'CONNELL, D. N. An investigation of cortical currents. *Proc. Amer. philos. Soc.*, 1952, **96**, 290-330.
6. KÖHLER, W., & WALLACH, H. Figural aftereffects: an investigation of visual processes. *Proc. Amer. philos. Soc.*, 1944, **88**, 269-357.
7. OSGOOD, C. E., & HEYER, A. W. A new interpretation of figural aftereffects. *Psychol. Rev.*, 1952, **59**, 98-118.
8. WERTHEIMER, M. The differential satiability of schizophrenic and normal subjects: a test of a deduction from the theory of figural aftereffects. *J. gen. Psychol.*, in press.

(Received January 4, 1954)

THE BRITISH JOURNAL OF PSYCHOLOGY

Edited by D. W. HARDING

Vol. XLV. Part 1 February 1954 12s. 6d. net.

L. S. HEARNshaw. The psychological study of conceptual thinking.

HILARY M. CLAY. Changes of performance with age on similar tasks of varying complexity.

R. C. OLDFIELD. Memory mechanisms and the theory of schemata.

C. B. GIBBS. The continuous regulation of skilled response by kinaesthetic feed back.

D. SHEPPARD. The adequacy of everyday quantitative expressions as measurements of qualities.

E. C. POULTON. Verbal extrapolation and interpolation.

M. B. SHAPIRO. A preliminary investigation of the effects of continuous stimulation on the perception of 'apparent motion'.

CORRESPONDENCE.

PUBLICATIONS RECENTLY RECEIVED.

Vol. XLV. Part 2 May 1954 12s. 6d. net.

F. V. SMITH. Critical notice and appreciation of the work of the late Professor Clark L. Hull.

SYLVIA ANTHONY. The effect of the vocational aims of industrial apprentices upon their attitude to education and commissioning.

GURTH HIGGIN. The effect of reference group functions on social status ratings.

FRIEDA GOLDMAN-EISLER. On the variability of the speed of talking and on its relation to the length of utterances in conversations.

JOHN COHEN, C. E. M. HANSEL and JOHN SYLVESTER. An experimental study of comparative judgments of time.

R. W. PICKFORD. Some problems of anomalous colour vision.

LESTER LUBORSKY. A note on Eysenck's article "The effects of psychotherapy: an evaluation."

H. J. EYSENCK. A reply to Luborsky's note.

CORRESPONDENCE.

PUBLICATIONS RECENTLY RECEIVED.

The subscription price per volume, payable in advance,
is 40s. net (post free). (U. S. \$6.50).

Subscriptions may be sent to any bookseller or to the

CAMBRIDGE UNIVERSITY PRESS

Bentley House, Euston Road, London, N. W. 1

PSYCHOLOGICAL REVIEW

July 1, 1954

YEAR	VOLUME	AVAILABLE NUMBERS					PRICE PER NUMBER	PRICE PER VOLUME
1894	1	-	2	-	4	5	\$1.50	—
1895	2	-	-	3	4	5	\$1.50	—
1896	3	-	-	-	-	-	—	—
1897	4	1	-	-	-	-	\$1.50	—
1898	5	-	-	3	-	5	\$1.50	—
1899	6	-	-	-	-	-	\$1.50	—
1900	7	1	-	-	-	-	\$1.50	—
1901	8	1	2	-	-	-	\$1.50	—
1902	9	-	2	-	-	-	\$1.50	—
1903	10	1	2	-	-	-	\$1.50	—
1904	11	1	-	-	4	5	\$1.50	—
1905	12	1	2	3	4	5	\$1.50	—
1906	13	-	-	3	4	5	\$1.50	—
1907	14	1	2	-	-	-	\$1.50	—
1908	15	-	-	-	-	-	—	—
1909	16	1	-	3	4	5	\$1.50	—
1910	17	1	2	3	-	-	\$1.50	—
1911	18	1	2	3	4	5	\$1.50	—
1912	19	1	2	3	4	5	\$1.50	\$8.00
1913	20	1	2	3	4	5	\$1.50	\$8.00
1914	21	-	2	3	4	5	\$1.50	—
1915	22	1	2	3	4	5	\$1.50	\$8.00
1916	23	1	-	-	4	5	\$1.50	—
1917	24	-	-	-	4	-	\$1.50	—
1918	25	-	2	3	4	5	\$1.50	—
1919	26	1	2	3	4	5	\$1.50	\$8.00
1920	27	1	-	-	-	-	\$1.50	—
1921	28	-	2	-	-	-	\$1.50	—
1922	29	1	-	-	4	-	\$1.50	—
1923	30	1	-	3	-	-	\$1.50	—
1924	31	1	2	3	4	5	\$1.50	\$8.00
1925	32	-	2	3	-	5	\$1.50	—
1926	33	1	-	3	4	5	\$1.50	—
1927	34	1	2	3	4	5	\$1.50	\$8.00
1928	35	1	2	3	4	5	\$1.50	\$8.00
1929	36	1	2	3	4	5	\$1.50	\$8.00
1930	37	1	2	3	4	5	\$1.50	\$8.00
1931	38	1	2	3	4	5	\$1.50	\$8.00
1932	39	1	2	3	4	5	\$1.50	\$8.00
1933	40	1	2	3	4	5	\$1.50	\$8.00
1934	41	1	2	3	4	5	\$1.50	\$8.00
1935	42	1	2	3	4	5	\$1.50	\$8.00
1936	43	1	2	3	4	5	\$1.50	\$8.00
1937	44	1	2	3	4	5	\$1.50	\$8.00
1938	45	1	2	3	4	5	\$1.50	\$8.00
1939	46	1	2	3	4	5	\$1.50	\$8.00
1940	47	-	-	3	4	-	\$1.50	—
1941	48	-	-	-	-	6	\$1.50	—
1942	49	-	2	3	4	5	\$1.50	—
1943	50	1	2	3	4	5	\$1.50	\$8.00
1944	51	1	2	3	4	5	\$1.50	\$8.00
1945	52	1	2	3	4	-	\$1.50	—
1946	53	1	2	3	4	5	\$1.50	\$8.00
1947	54	1	2	-	4	5	\$1.50	—
1948	55	1	2	3	4	5	\$1.50	\$8.00
1949	56	1	2	3	4	5	\$1.50	\$8.00
1950	57	1	2	3	4	5	\$1.50	\$8.00
1951	58	1	2	3	4	5	\$1.50	\$8.00
1952	59	1	2	3	4	5	\$1.50	\$8.00
1953	60	1	2	3	4	5	\$1.50	\$8.00
1954	61	By subscription \$6.50, foreign \$7.00					\$1.50	—

Postage prepaid on U. S. orders. Add \$.50 per volume on foreign orders. All stock subject to prior sale. The American Psychological Association gives the following discounts on orders for any one journal:

10% on orders of \$ 50.00 and over
20% on orders of \$100.00 and over
30% on orders of \$150.00 and over

Current subscriptions and orders for back numbers should be addressed to

AMERICAN PSYCHOLOGICAL ASSOCIATION
1333 Sixteenth Street N.W. Washington 6, D. C.